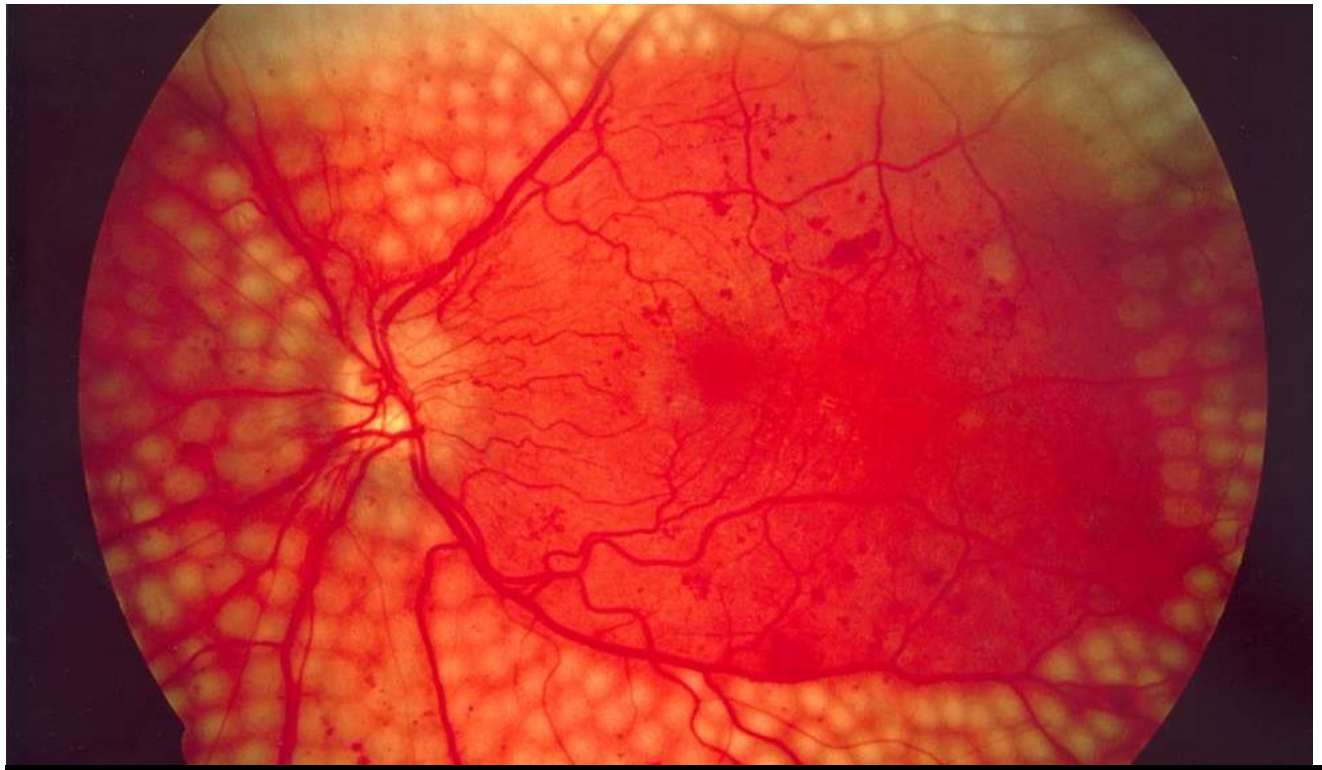


Neutrophil lymphocyte ratio as a Predictor of diabetic retinopathy Incidence in Scottish population

- **Aravind Lathika Rajendrakumar**

  PhD Candidate

- **Prof. Colin Palmer**

  Supervisor

FUNDED BY NIHR | National Institute for Health Research

Presentation Outline

- Introduction on Diabetic retinopathy (DR) and Neutrophil to Lymphocyte Ratio (NLR)
- Data source description
- Competing risks analysis of NLR with DR
- Summary of the findings
- Association of NLR with biomarkers in GoDARTS cohort
- GWAS of NLR in the Scottish population
- Protein-Protein Interaction Network and Gene-Gene Interaction Network

Data Source

- Tayside & Fife
- Genetics of Diabetes Audit and Research in Tayside Scotland (GoDARTS)
- Genetics of Scotland Health Research Register (GoSHARE)
- Madras Diabetes Research Foundation (MDRF)

Introduction

Diabetic Retinopathy

- microvascular complication with close to 35% of diabetic patients reporting DR globally
- Heterogeneity in DR is yet to be fully uncovered
- Explicit role of the immune system in DR risk has not been explored in detail
- Deaths mask the development of DR in individuals with diabetes

Neutrophil to Lymphocyte Ratio (NLR)

- white blood cell components that play an important role in immunity
- composite marker of inflammation which is routinely available as a part of clinical investigations
- more robust to variations and provide more predictive information than its component markers
- represents subclinical inflammation – a prominent feature reported in chronic diseases and is generally high in individuals with diabetes.
- Neutrophils secrete different inflammatory molecules that affect the integrity of blood vessels whereas lymphocytes act as are more of modulators of inflammatory activity

Aims and objectives

- To describe the secular trends of DR and longitudinally examine the association between NLR and diabetic retinopathy under a competing risks model
- To identify the genetic locus of NLR and its biomarker associations using the Scottish genetic data
- To identify specific gene-gene and protein-protein interactions from the above analysis using network visualization using established datasets

Phenotype definition

- Incidence of DR was defined as the time to the first diagnosis of R1 or above grade in the Scottish retinopathy grading scheme from diabetes diagnosis
- NLR was defined as the ratio of absolute counts of Neutrophils divided by absolute count of Lymphocytes

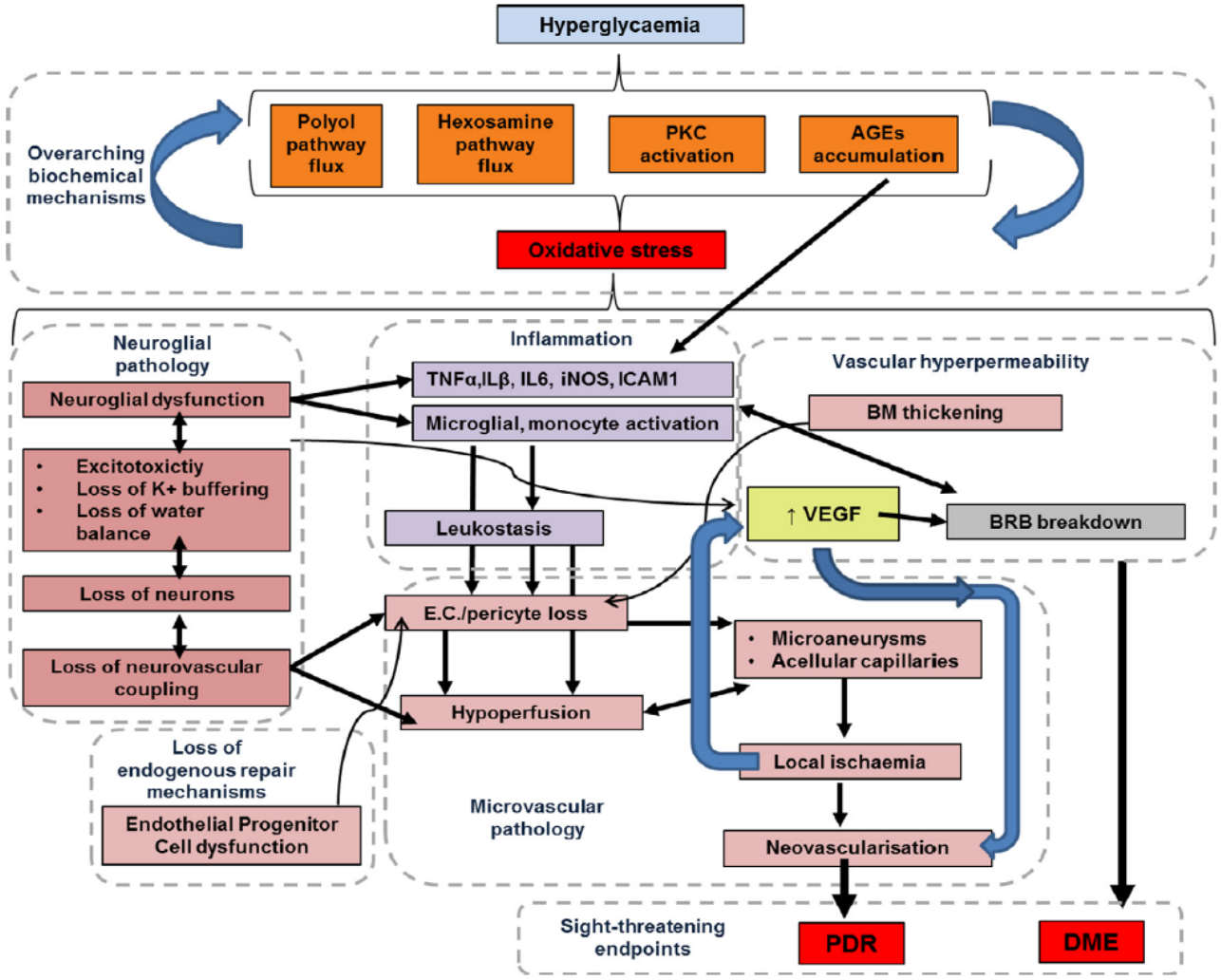
Phenotype definition (continued)

- Diabetic maculopathy status was not considered
- The grades range from 0-4 which indicates increasing severity of retinopathy from no retinopathy to severe proliferative retinopathy.
- Prior to the analysis, it was decided that all records with laser photocoagulation would be marked as R4 and for incidence analysis
- Participants who already had retinopathy at baseline would be excluded.
- Eye having a severe grade was considered
- Excluded all the individuals with NLR > 20 at baseline

Statistical Analysis

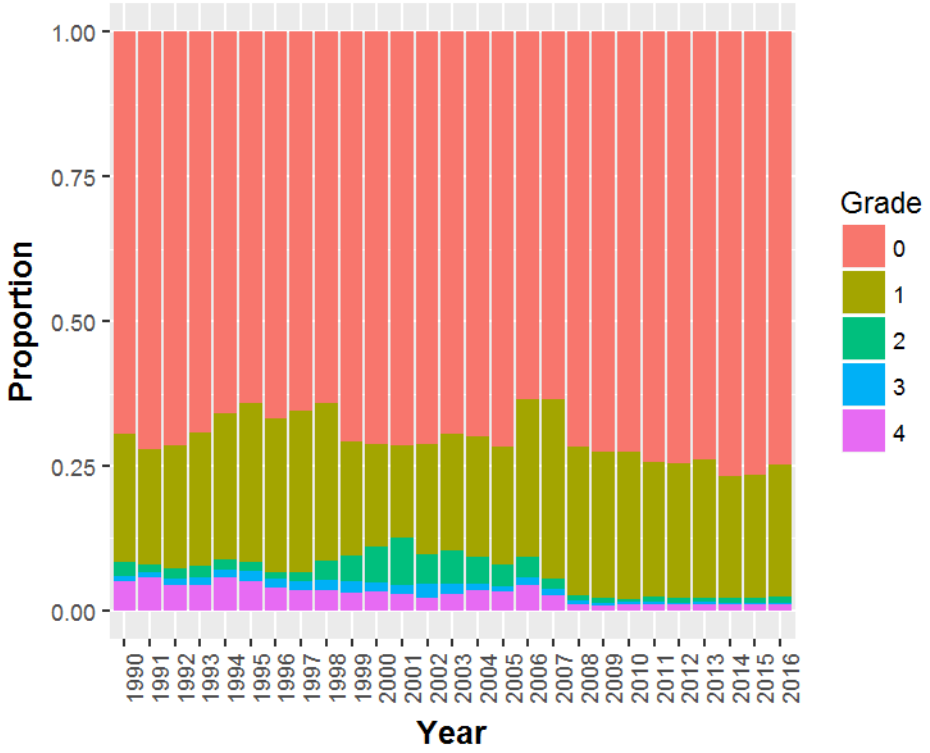
- The effect of NLR and its interactions were explored using a competing risks model
- Fine-Gray model to predict the effect of NLR

Biochemical mechanisms involved in the pathogenesis of Diabetic Retinopathy



Source: Lechner J, O'Leary OE, Stitt AW. The pathology associated with diabetic retinopathy. Vision Res;139:7–14.

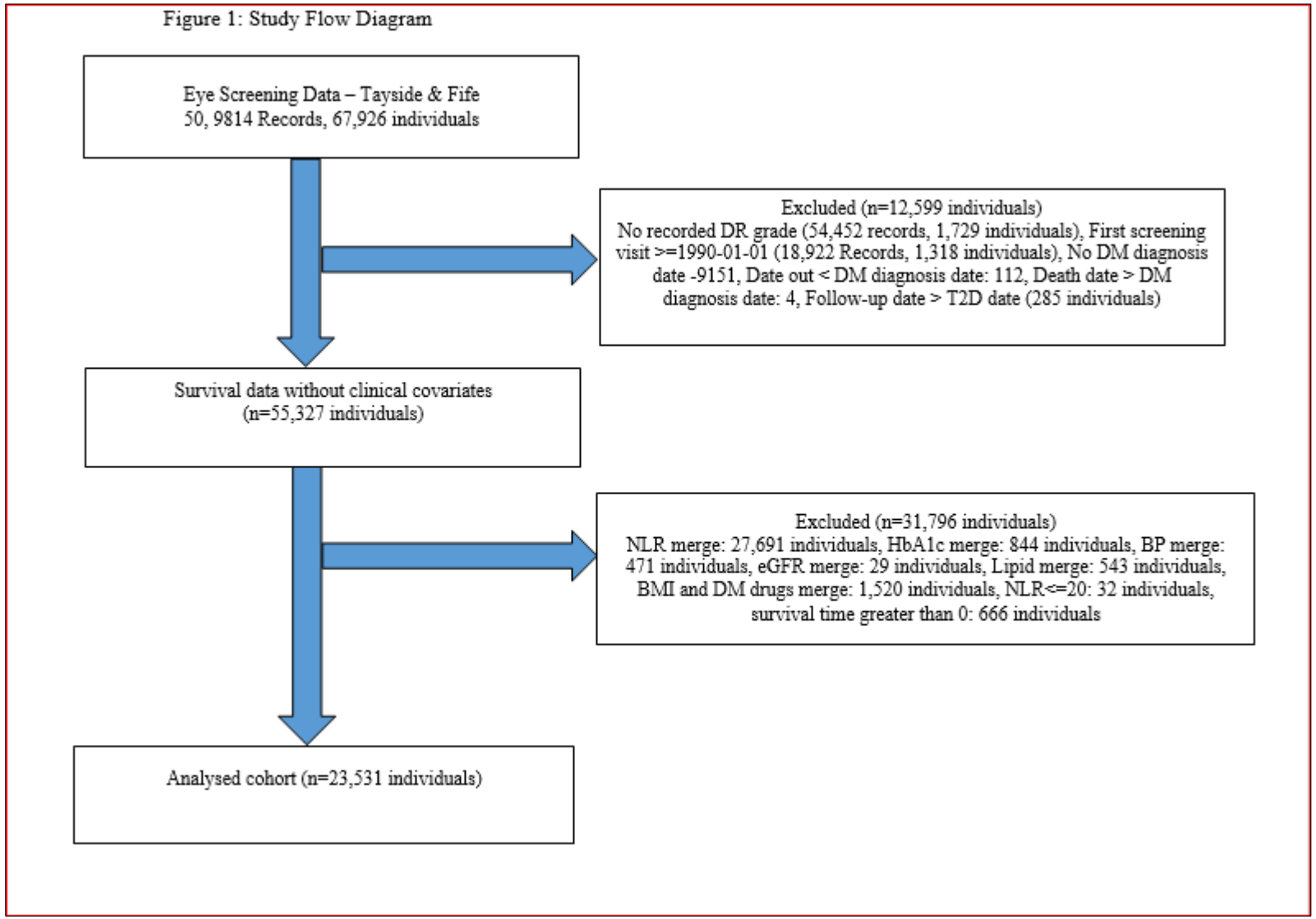
Trend in the prevalence of DR in Tayside and Fife from 1990-2016*



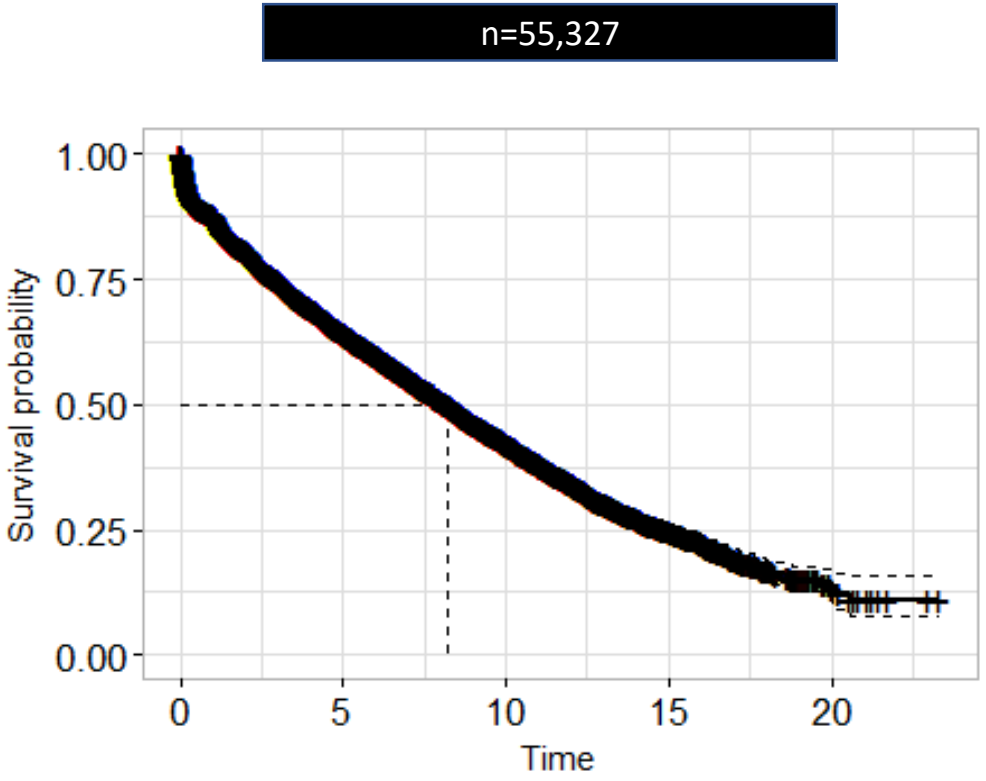
n=64,879

Subject may appear more than once in the plot*

Study Flow Diagram

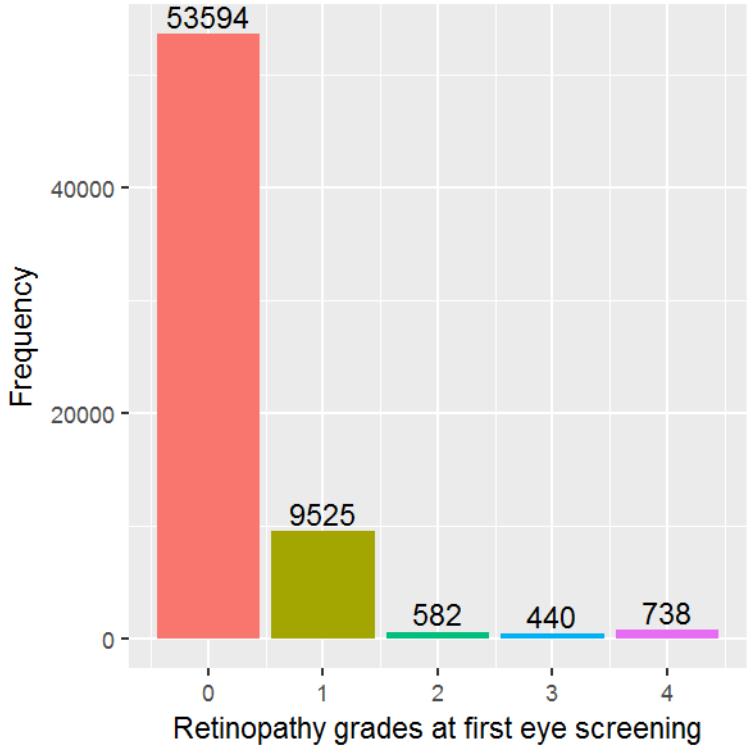


Survival analysis – Tayside & Fife screening data set



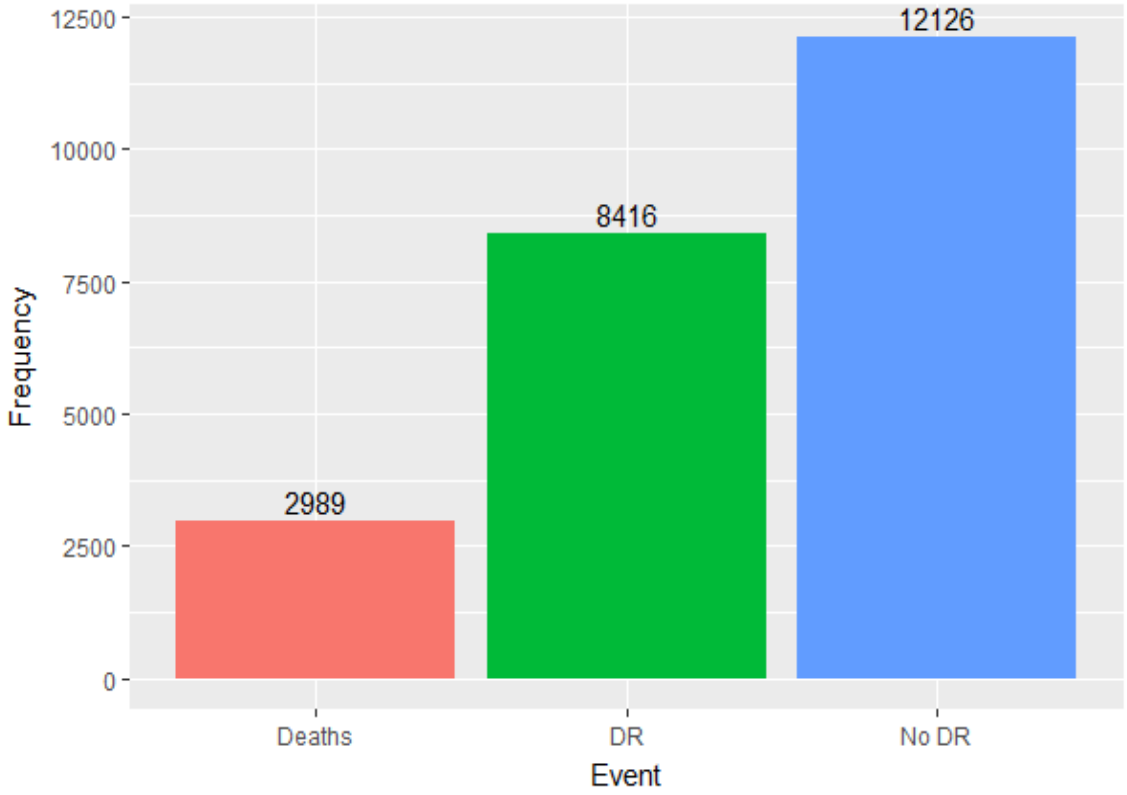
Incident DR =26214,Median -9.55 yrs,0.95 LCL-9.42,0.95 UCL-9.68

Prevalence of DR at the first retinal screening in Tayside and (n=64,879)

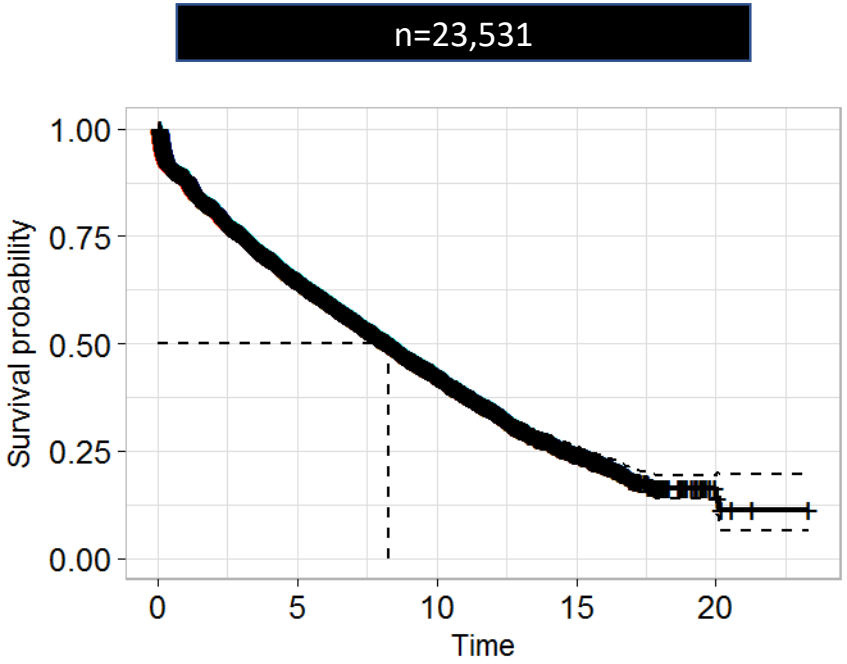


17.4% had DR at baseline, n=64,879

Events at the end of 10 Years follow up after Diabetes diagnosis in Tayside and Fife (n=23,531)



Complete Case Analysis- DR (10 Year Follow UP)



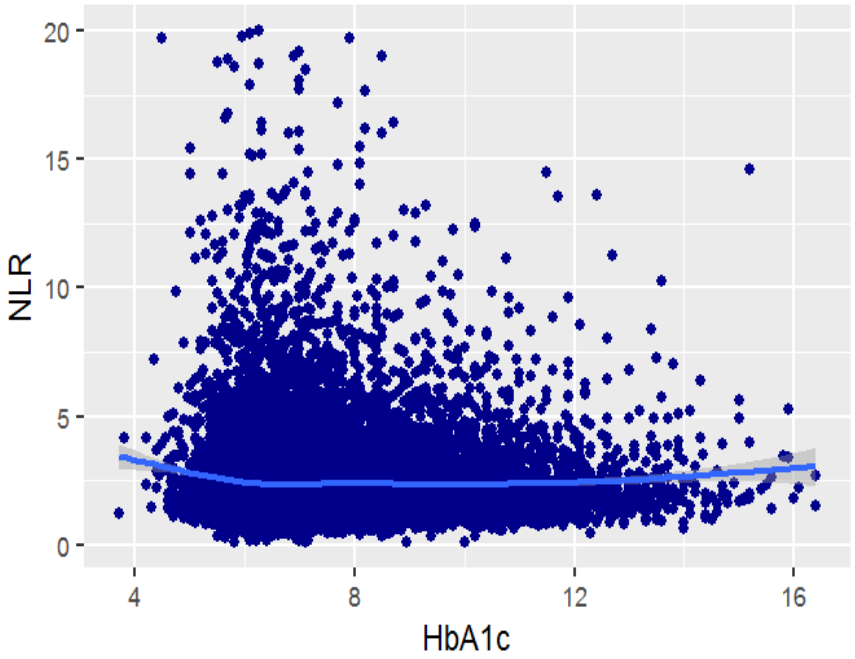
median follow up time was 3.3 years

Incident DR =8876, Median -8.27 yrs., 0.95 LCL-8.08 yrs., 0.95 UCL-8.48 yrs.

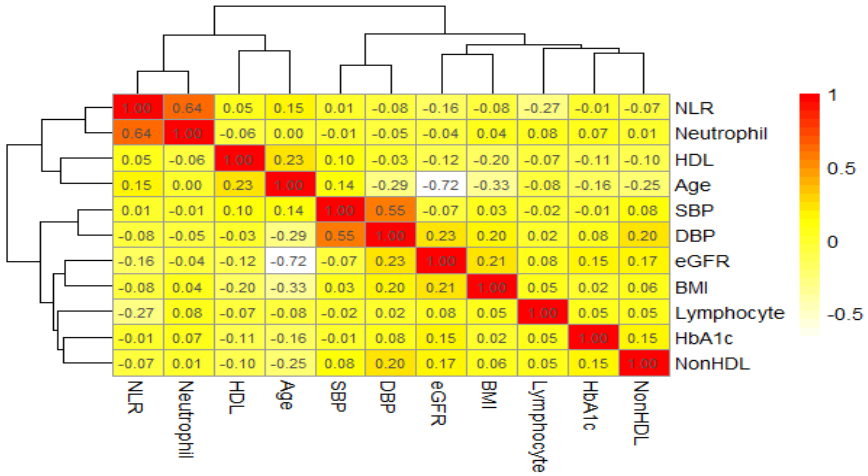
Demographic and clinical characteristics of participants at baseline (n=23,531)

Parameter	Mean	SD	Range
Age (Years)	61.7	12.7	17.1-96.5
Male (%)	55.3%		
Diabetes Drug (Yes %)	45.7%		
HbA1c (%)	7.3	1.5	3.7-16.4
SBP(mmHg)	140.3	17.3	72-240
DBP(mmHg)	81.3	9.9	40-142.5
eGFR (ml/min/1.73m^2)	80.6	19.6	15.5-163.2
Lymphocytes (10^9/L)	2.3	1.62	0.3-187.2
Neutrophils (10^9/L)	4.8	1.9	0.4-28.5
NLR	2.4	1.5	0.08-20.0
BMI (kg/m2)	32.2	6.6	15.2-73.9
HDL-c (mmol/L)	1.2	0.3	0.1-3.9
Non-HDL-c (mmol/L)	3.7	1.1	0.6-18.4

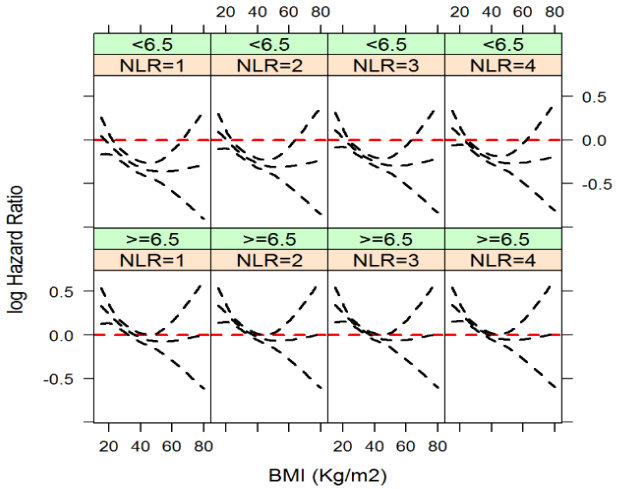
Correlation of NLR with HbA1c at baseline in Tayside cohort

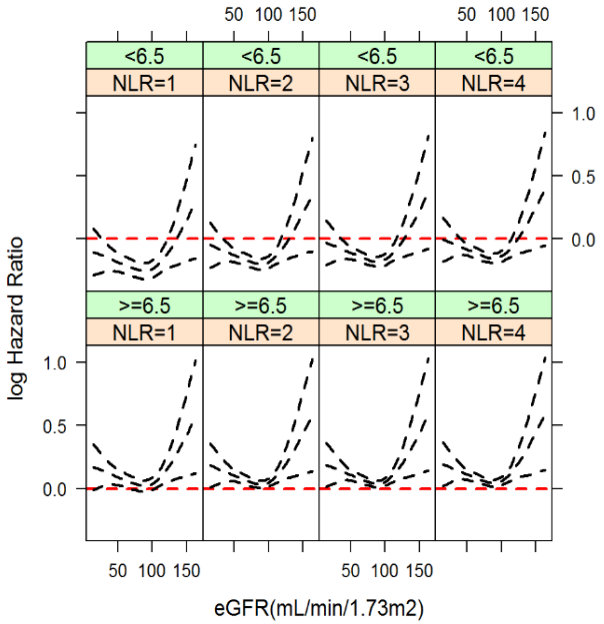


Correlation Heat Map

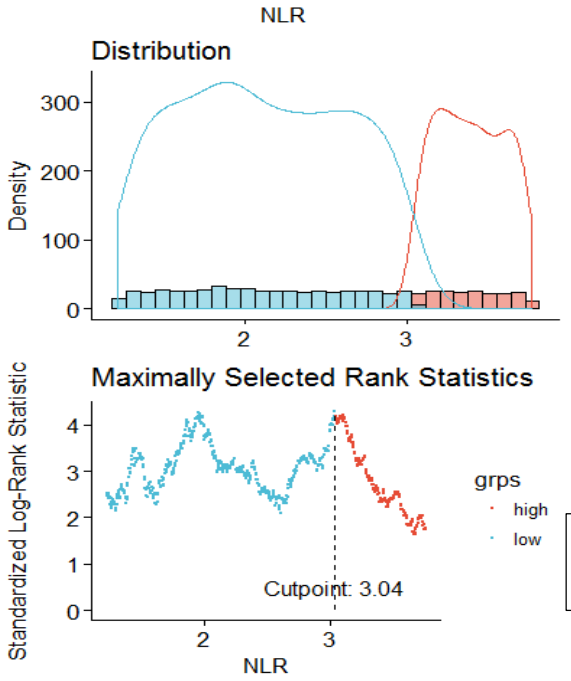


Three variable Conditional plot showing interaction depicting non –linear effects of (a) BMI (b) eGFR for glycaemic levels (high and low) in quartiles of NLR



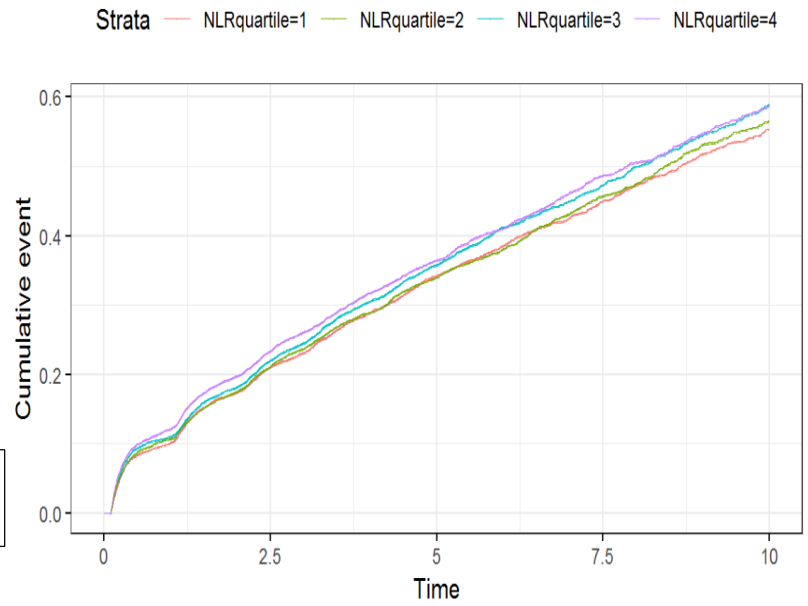


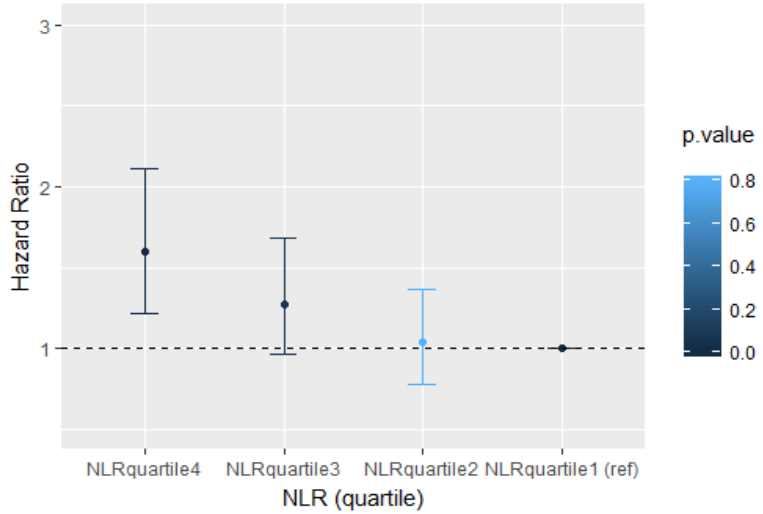
Optimal cut-point for NLR determined using maximal selected rank statistics



NLR > 3.04 was coded as high
4406 (18.7%) individuals had high (>3.04) NLR values

Cumulative incidence plot of hazards of DR incidence by NLR for a 10 year follow-up period in Tayside





Adjusted Hazards for incidence of DR for different quartiles of NLR from Cause-Specific hazard Model

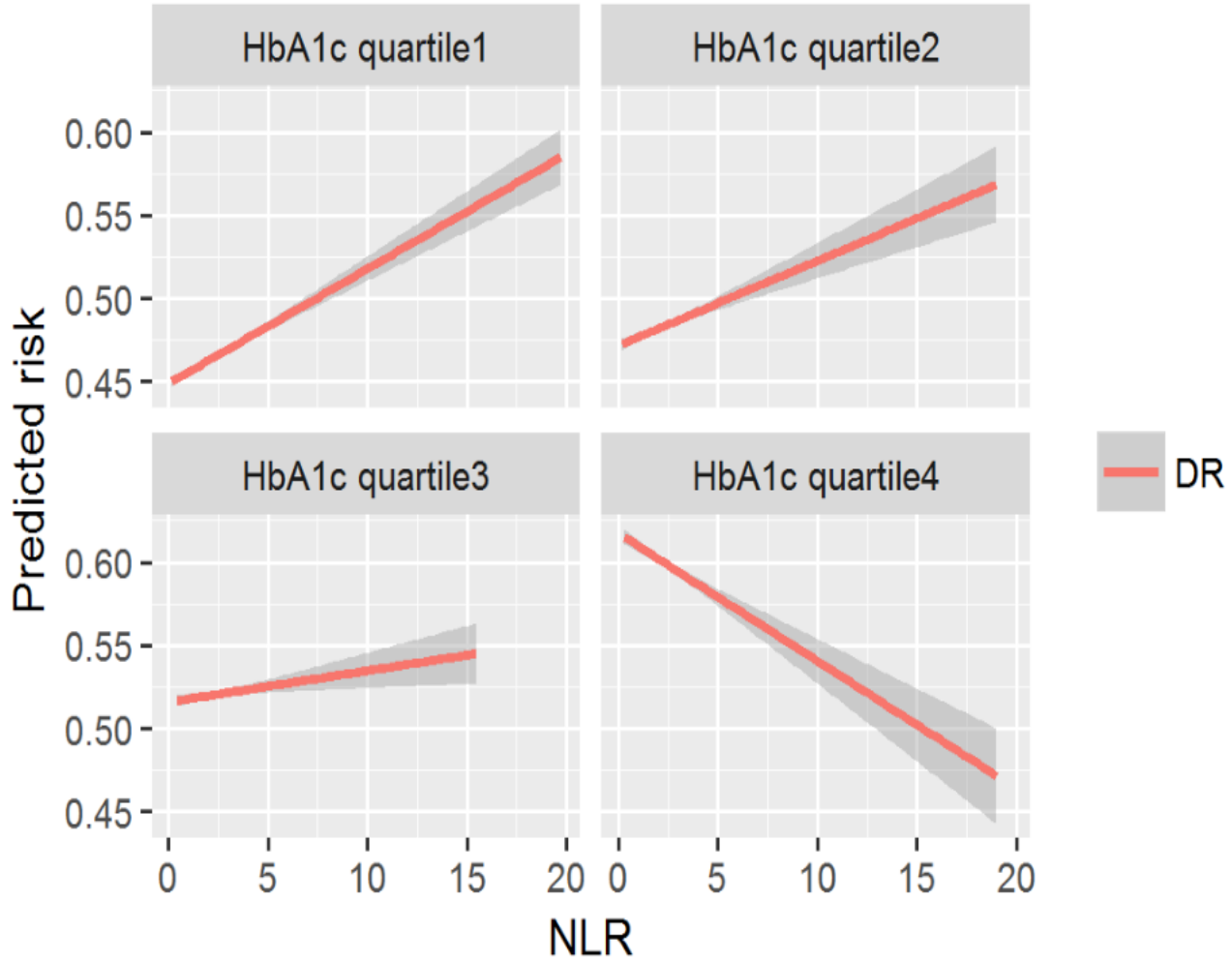
Results of Cause-Specific and Fine-Gray regression model for 10-year Incidence of DR (n=23,531)

Parameter	CSH (95% CI)	P	SH (95%CI)	P
Sex(M)	1.08(1.04,1.13)	<0.001	1.05 (1.01, 1.10)	<0.05
DBP(mm Hg)	0.99(0.99,0.99)	<0.05	-	-
HbA1c (%)	1.15(1.11,1.20)	<0.001	1.19(1.15, 1.24)	<0.001
HbA1cxNLR	0.94 (0.91,0.97)	<0.01	0.94 (0.90, 0.96)	<0.001
SBP(mmHg)	1.00(1.00,1.01)	<0.001	1.00(1.00, 1.00)	<0.001
Age x NLR	-	-	0.99(0.99, 0.99)	<0.001
NLR(>3.04)	1.63(1.28,2.07)	<0.001	2.24(1.70, 2.94)	<0.001
Non HDL-c(mmol/L)	0.96(0.94,0.98)	<0.001	0.98 (0.98, 0.98)	<0.01
BMI (kg/m2)	0.99(0.99,0.99)	<0.001	0.98(0.98, 0.98)	<0.001
eGFRml/min/1.73 m^2	-	-	0.99(0.99,0.99)	<0.001
Diabetes drug (Yes)	1.11(1.06,1.17)	<0.001	-	-
Adjusted for Age(years), HDL-c(mmol/L), eGFRxNLR				

Subgroup analysis of NLR showing the difference in predicted risk for DR in younger (<=45 years) and versus older (>=65 years) age groups with better glycaemic control (<=7.0 %)

Parameter	Age (<=45 years)	P	Age (>=65 years)	P
	(n=1,055)		(n=6,285)	
NLR(>3.04)	1.45 (1.08,1.94)	<0.05	1.10 (0.99,1.22)	0.059
Adjusted for Sex, SBP(mmHg), DBP(mm Hg), BMI (kg/m2), Non HDL-c(mmol/L),HDL-c(mmol/L), Diabetes drug , eGFRml/min/1.73 m^2				

Risk for DR predicted by NLR stratified by quartiles of HbA1c estimated from covariate-adjusted Fine-Gray model (Hba1c quartile 4 had the highest HbA1c)



Summary and discussion

- powerful relationship between NLR levels at diagnosis of diabetes and the risk of retinopathy over a 10 year period

- both standard and “competing risk” methodologies to confirm that this relationship is independent of overall effects on mortality

- increased NLR does not provide independent additional prediction in individuals of advanced age or with poor diabetes control, but its utility is limited to younger individuals with better glycaemic control

- possible unmeasured confounders resulting from the observational nature of the data may have affected our study results

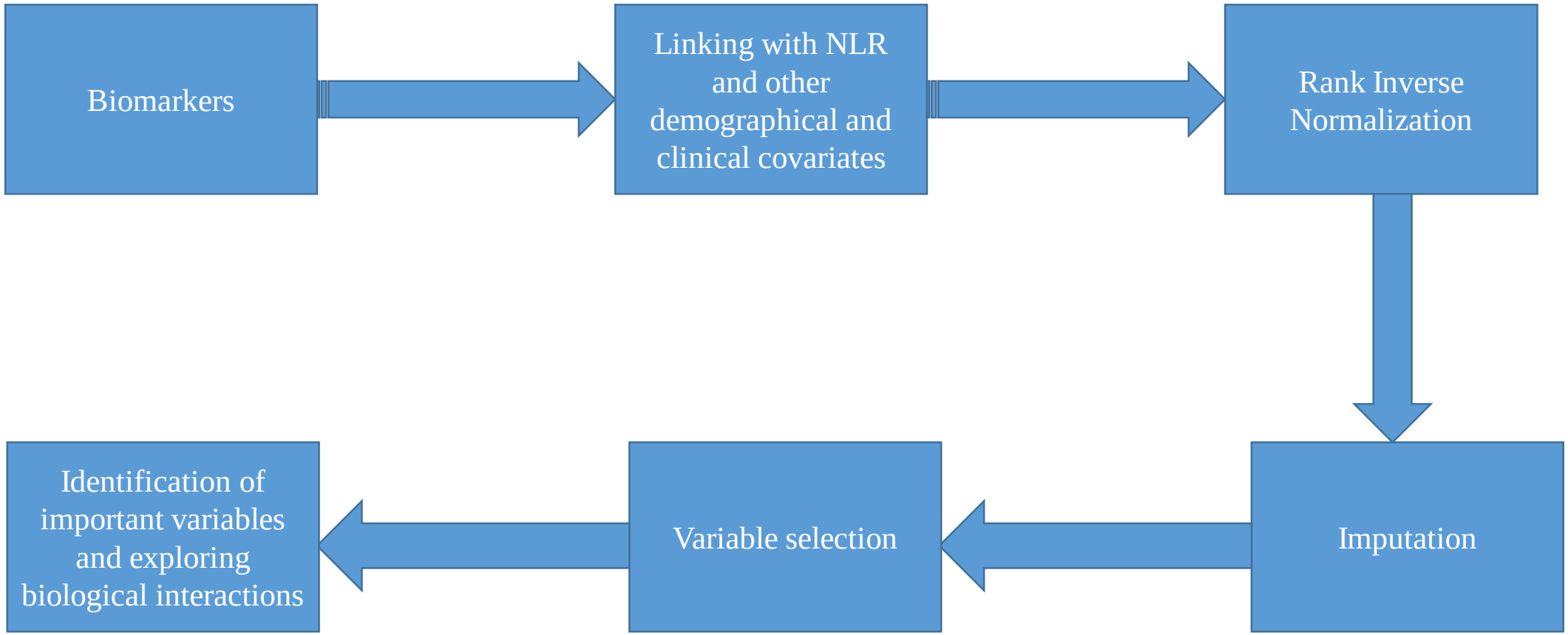
Association of NLR with Biomarkers from GoDARTS Cohort

What is RHAPSODY study all about?

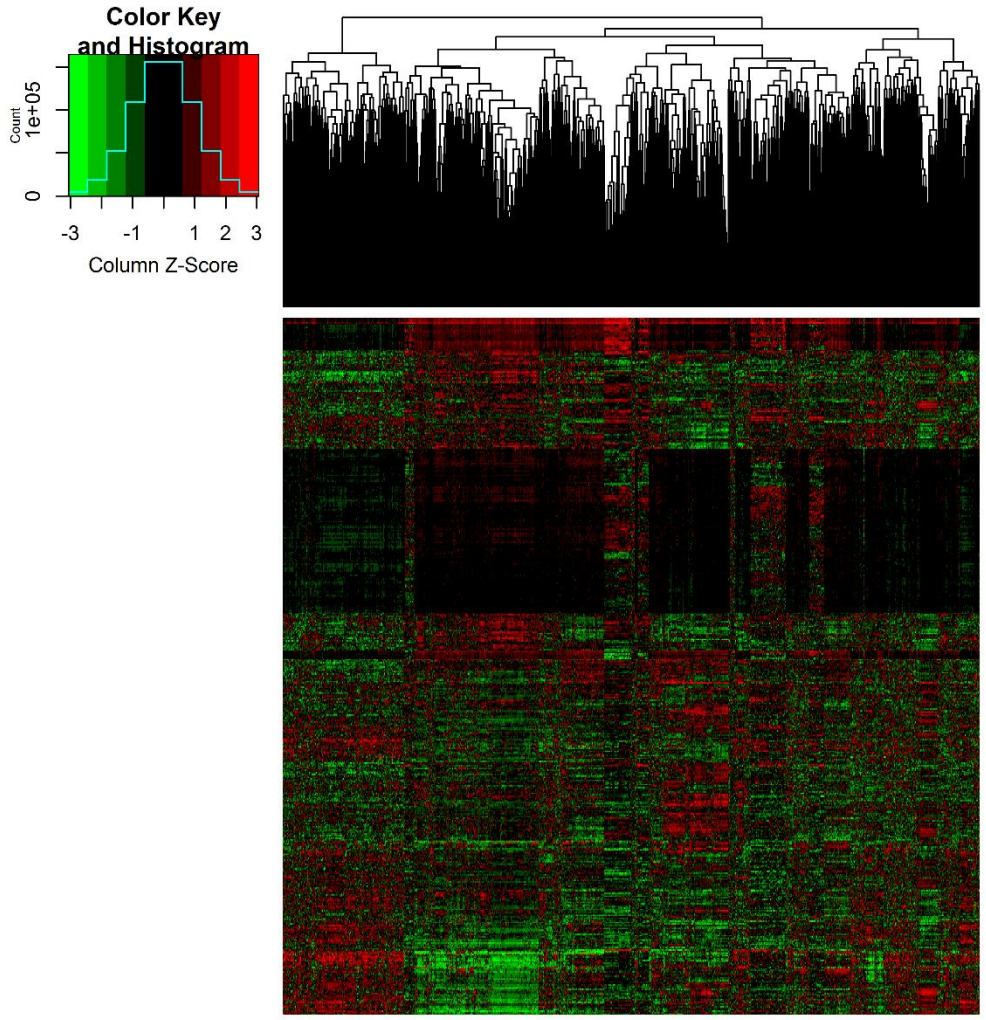
- Comprehensive multi-omics database
- Peptidomic, lipidomic and polar metabolites analytic platforms to measure biomarkers to predict the trajectory from borderline diabetes to full blown diabetes and subsequent progression.
- 10 cohorts from 5 European countries
- Dundee is a part of this consortium!!!
- We assayed 1417 markers … amazing is it not??!!!!
- Diabetic Kidney disease and diabetic retinopathy associations with GoDARTS cohort are being explored currently

Academia Université de Lausanne Switzerland
Lunds Universitet Sweden
Technische Universität Dresden Germany
Università di Pisa Italy Université Paris Diderot – Paris 7 France
INSERM France
Université Libre de Bruxelles Belgium
Institut Suisse de Bioinformatique Switzerland
University of Oxford United Kingdom
CNRS France University of Eastern Finland Finland
University of Dundee United Kingdom
Imperial College London United Kingdom
Eberhard Karls Universität Tübingen Germany
Kobenhavns Universitet Denmark University hospitals
A.O.U. Città della Salute e della Scienza di Torino Italy
UMCG – Groningen The Netherlands
CHRU Lille France LUMC – Leiden The Netherlands
VUmc – Amsterdam
The Netherlands Pharmaceutical Industries (EFPIA) Institut de Recherches Servier France
Janssen Pharmaceutica NV Belgium Novo Nordisk A/S Denmark Sanofi-Aventis Deutschland GMBH Germany
Small and Medium-sized Entreprises SCIPROM Sàrl Switzerland
Lipotype GmbH Germany

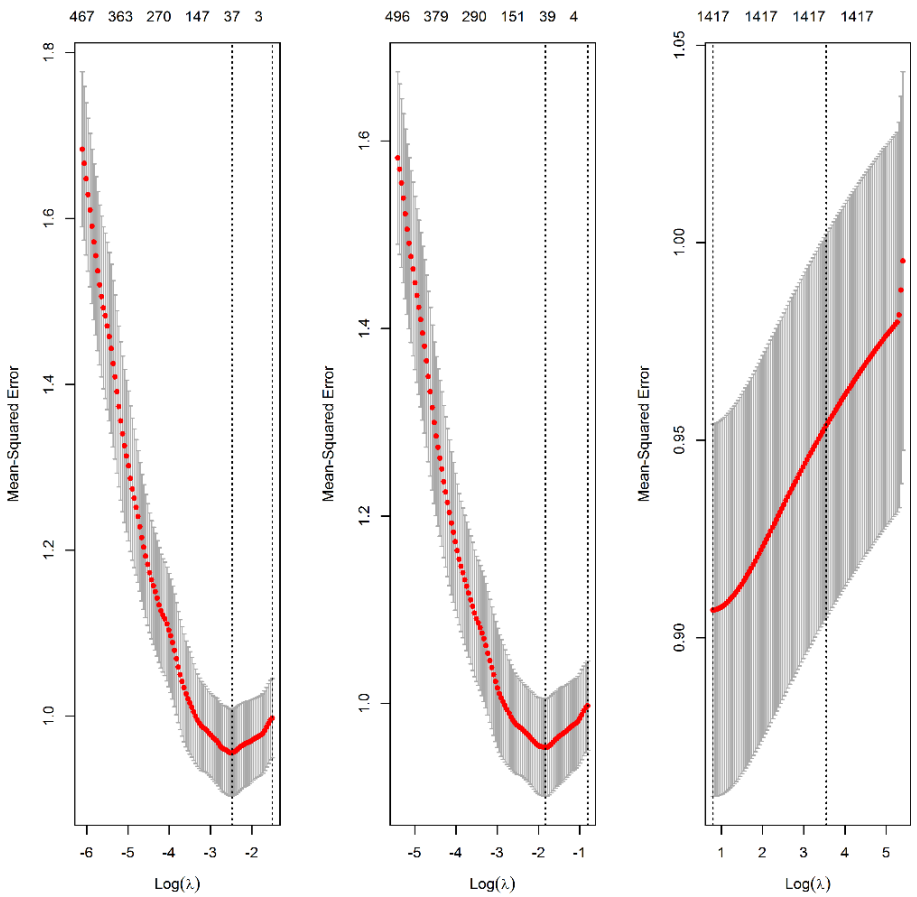
Analysis Flow Diagram

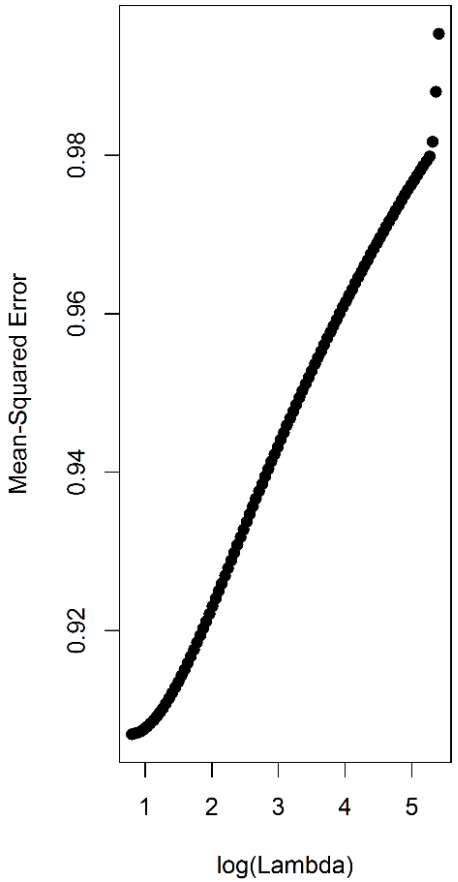


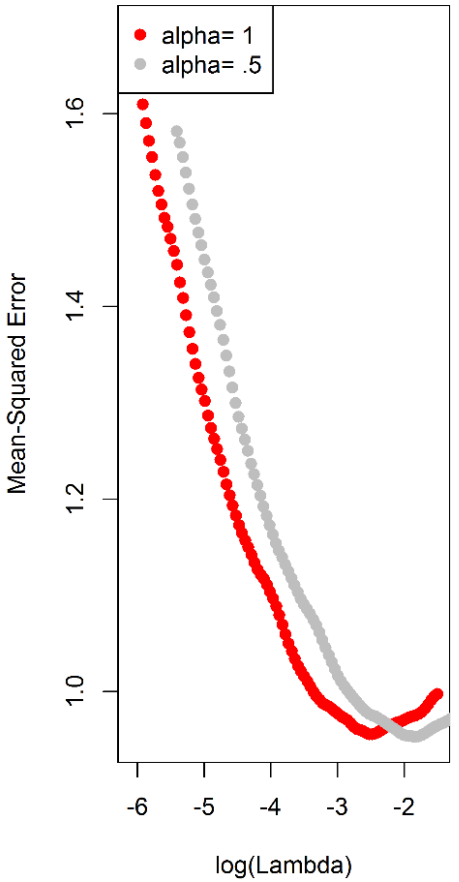
Heat map showing RHAPSODY Biomarker Correlations



Plots indicating selection of tuning parameters of alpha from the test data





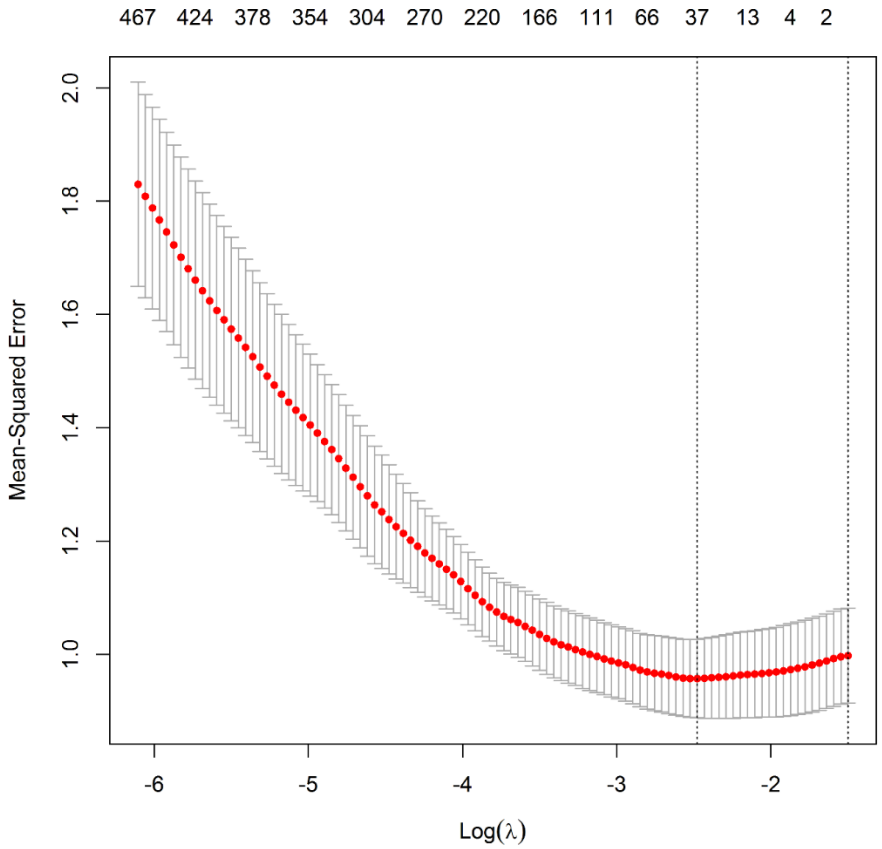


Summary statistics from full dataset

- 5 fold cross validation
- n= 480
- based on MSE
- alpha=1 and minimum λ
- Λ min – lambda corresponding to minimum MSE in the model
- 37 variables with non – zero coefficients selected by the model

Alpha=1	lambda	Measure	SE
min	0.083	0.957	0.069
1se	0.223	0.997	0.083

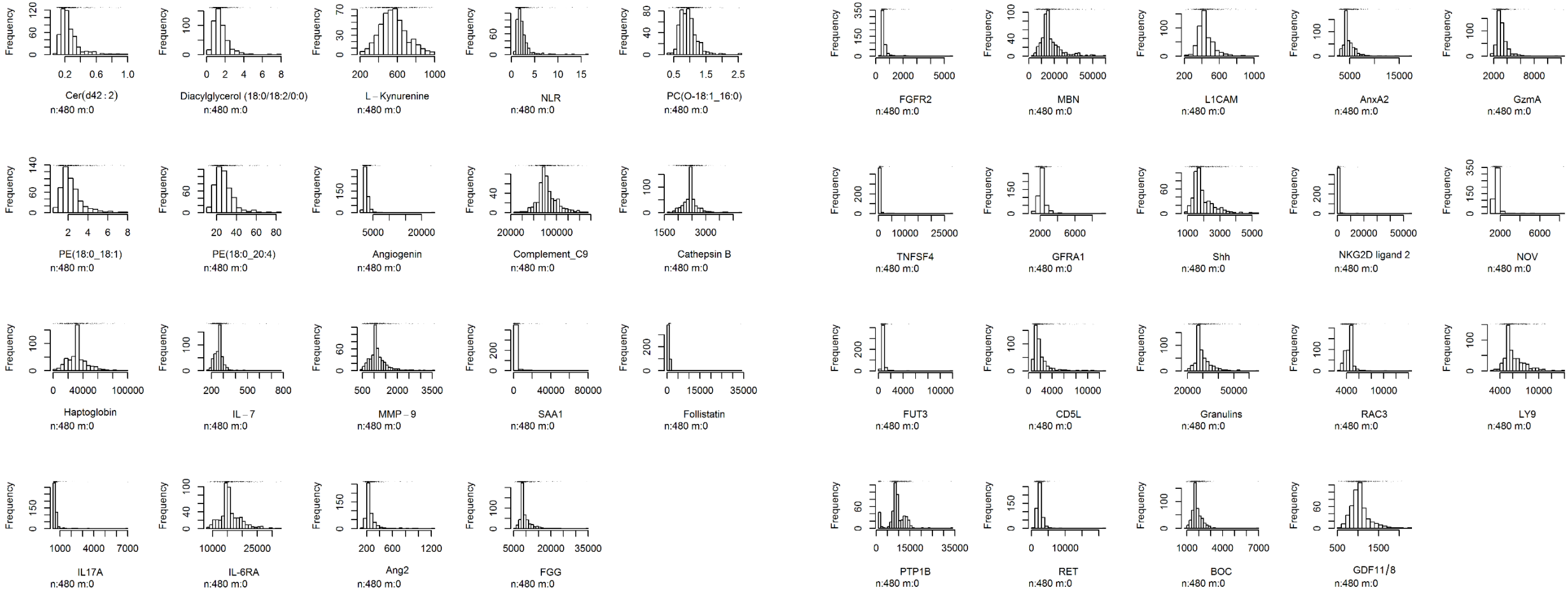
Plot showing fit with the optimal model parameters



Biomarkers associated with NLR identified by elastic net

Code	Biomarker	Coefficients
Kynu	L-kynurenine	0.022688
SL000003	Angiogenin	0.003033
SL000325	Complement component C9	0.014925
SL000343	Cathepsin B	0.006689
SL000437	Haptoglobin	0.01366
SL000483	Interleukin-7	0.021571
SL000527	Matrix metalloproteinase-9	0.016925
SL000572	Serum amyloid A-1 protein	0.016943
SL000674	Follistatin	0.018976
SL001713	Interleukin-17A	0.002932
SL001943	Interleukin-6 receptor subunit alpha	0.003097
SL001996	Angiopoietin-2	0.041601
SL003341	Fibrinogen gamma chain	0.018979
SL003990	Fibroblast growth factor receptor 2	0.009232
SL004008	Myeloblastin	0.044209
SL004154	Neural cell adhesion molecule L1	-0.03188
SL004209	Annexin A2	-0.00709
SL004298	Granzyme A	-0.01025
SL004649	Tumor necrosis factor ligand superfamily member 4	0.003996
SL004858	GDNF family receptor alpha-1	-0.00357
SL005220	Sonic hedgehog protein	-0.00479
SL005228	NKG2D ligand 2	0.01731
SL005236	Protein NOV homolog	0.018292
SL005575	Galactoside 3(4)-L-fucosyltransferase	-0.00511
SL006108	CD5 antigen-like	0.018737
SL007173	Granulins	0.021398
SL007310	Ras-related C3 botulinum toxin substrate 3	-0.0371
SL007674	T-lymphocyte surface antigen Ly-9	-0.0272
SL008967	Tyrosine-protein phosphatase non-receptor type substrate 1	-0.02455
SL010378	Proto-oncogene tyrosine-protein kinase receptor Ret	-0.05613
SL013490	Brother of CDO	0.036285
SL021043	Growth/differentiation factor 11/8	-0.00805
Cer_42_0_2	Ceramide (d42:0)	-0.09655
DAG_18_0_0_18_2_0	Diacylglycerol (18:0_18:2)	-0.03777
PC_O_18_1_0_16_0_0	Phosphatidylcholine (O-18:1_16:0)	0.015224
PI_18_0_0_18_1_0	Phosphatidylethanolamine (18:0_18:1)	-0.00068
PI_18_0_0_20_4_0	Phosphatidylethanolamine (18:0_20:4)	-0.00655

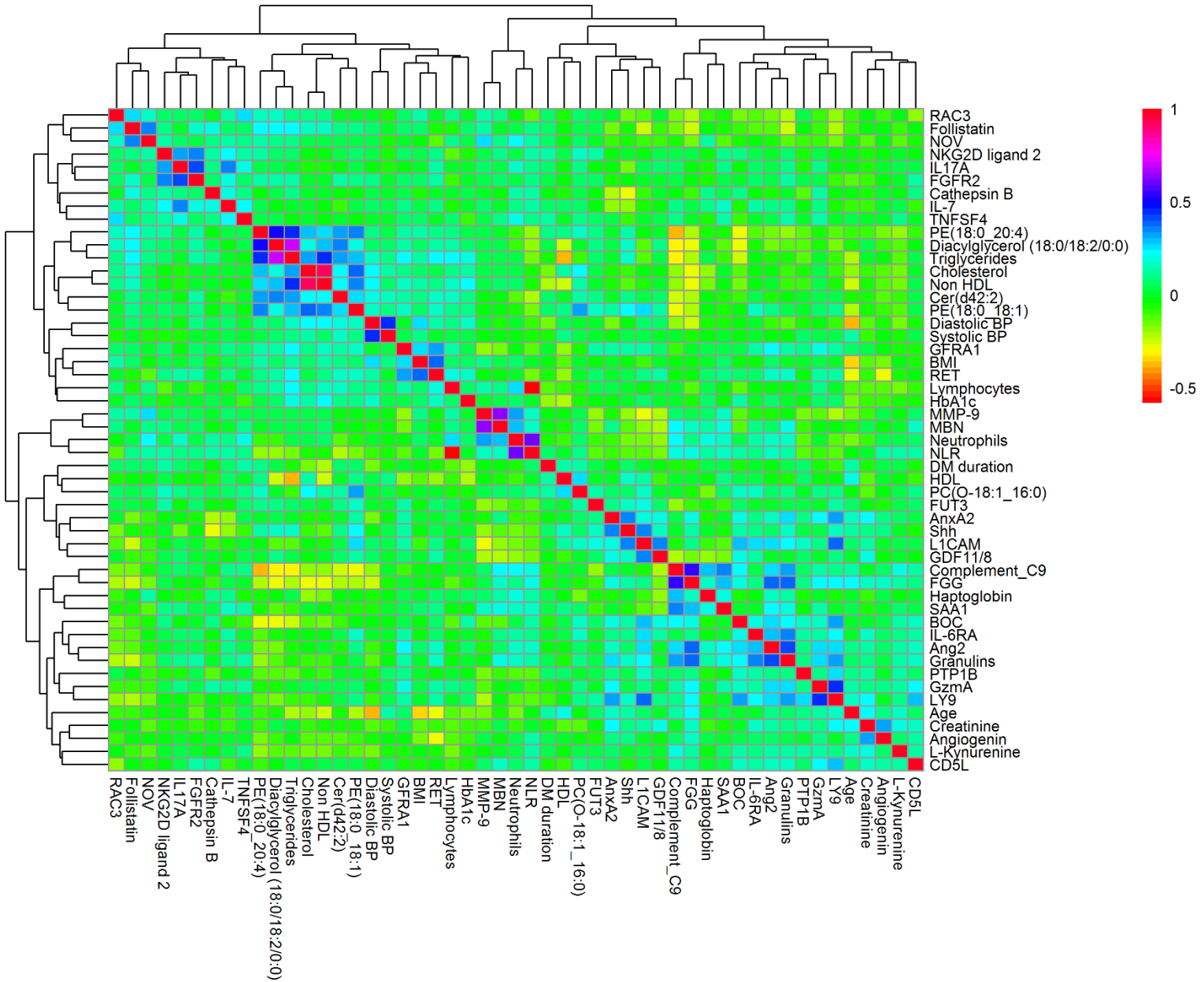
Biomarker distribution



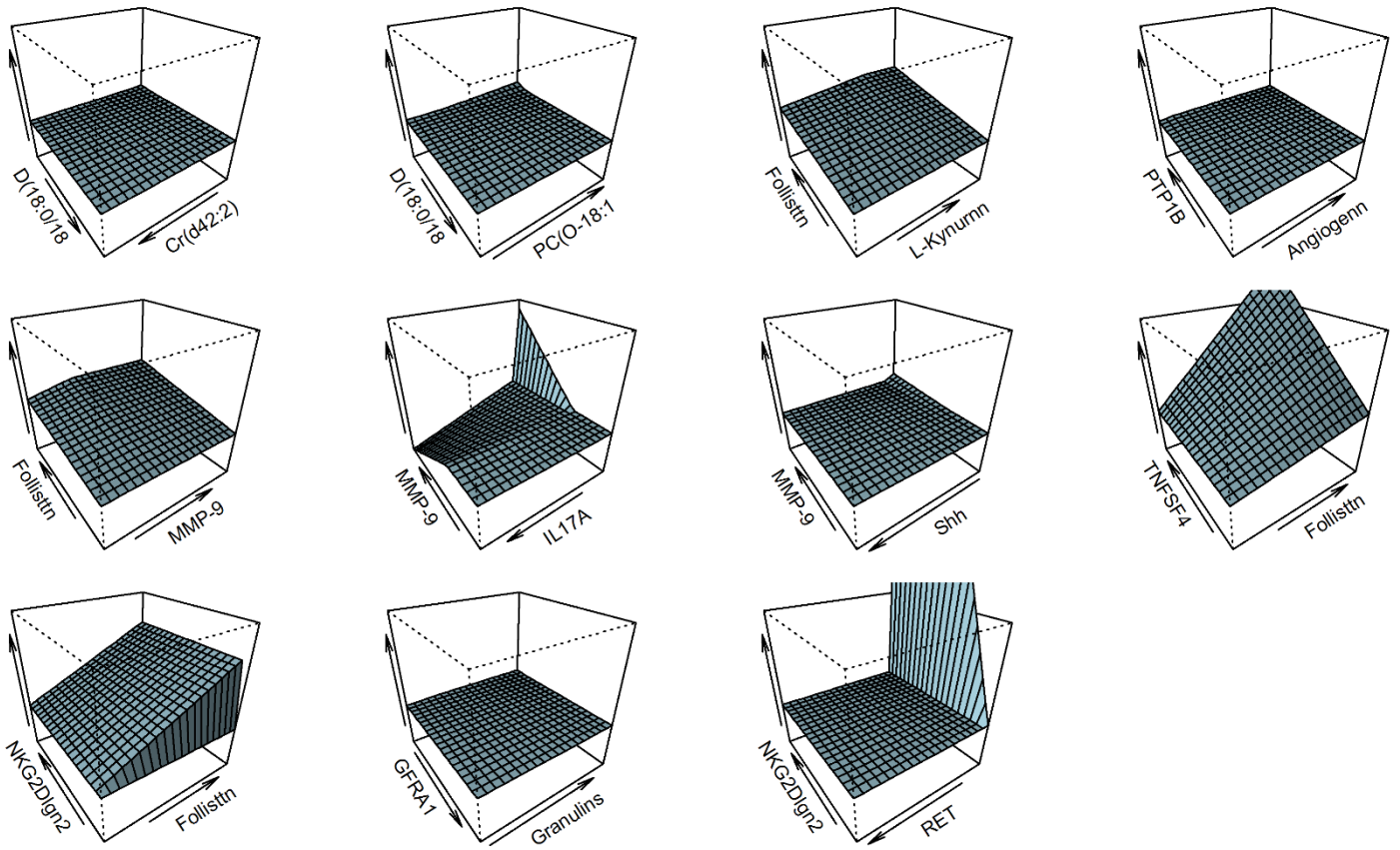
Summary statistics of all 37 biomarkers selected from LASSO regression

Biomarker	n	mean	sd	median	mad	min	max	range	skew	kurtosis	se
Cer(d42:2)	480	0.25	0.12	0.22	0.07	0.087461	0.98	0.89	2.36	7.68	0.005
Diacylglycerol (18:0/18:2/0:0)	480	1.48	0.80	1.34	0.59	0.271381	7.60	7.33	2.39	11.55	0.03
L-Kynurenine	480	558.69	147.21	548.53	132.77	203.7969	984.42	780.62	0.37	-0.0045	6.719
NLR	480	2.32	1.46	2	0.77	0.76	16.2	15.44	4.30	28.15	0.06
PC(O-18:1_16:0)	480	0.945594	0.26	0.91	0.21	0.367609	2.54	2.180	1.59	5.68	0.01
PE(18:0_18:1)	480	2.321292	1.03	2.06	0.75	0.603524	7.67	7.067	1.63	3.79	0.04
PE(18:0_20:4)	480	27.45	9.39	25.83	7.18	11.64146	83.88	72.24	1.74	5.39	0.42
Angiogenin	480	2955.78	1125.98	2766.782	360.64	1288.9	23041.1	21752.2	12.13	208.96	51.39
Complement_C9	480	83677.1	17616.07	80606.96	11908.1	27368	152128.6	124760.6	0.65	1.44	804.05
Cathepsin B	480	2443.20	284.30	2452.16	152.71	1631.3	4264	2632.7	1.28	6.36	12.97
Haptoglobin	480	32099.91	13077.36	31245.23	9186.19	165.5	95759.1	95593.6	0.84	2.12	596.89
IL-7	480	265.20	50.67	267.430	27.25	165.9	763.1	597.2	4.14	31.71	2.31
MMP-9	480	1143.46	383.78	1078.59	247.73	420.7	3570.2	3149.5	1.90	6.68	17.51
SAA1	480	1944.55	5438.57	651.61	373.58	176.3	78745.2	78568.9	9.08	104.82	248.23
Follistatin	480	1024.86	1503.43	950.03	63.15	585	33803	33218	21.58	467.71	68.62
IL17A	480	627.44	372.58	578.59	52.85	411.1	6995.7	6584.6	12.12	185.53	17.00
IL-6RA	480	15929.08	3562.49	15124.73	2085.20	8257.9	32802.1	24544.2	1.12	2.51	162.60
Ang2	480	258.04	99.00	227.539	32.02	137.8	1206.4	1068.6	4.61	30.15	4.51
FGG	480	9619.14	2759.69	8686.99	978.50	5479.7	34657.1	29177.4	3.37	18.67	125.96
FGFR2	480	617.359	328.20	546.26	64.64	387.7	5578.6	5190.9	8.82	112.86	14.98
MBN	480	16782.93	7990.55	14452.59	4454.69	3257.7	58345.3	55087.6	1.88	4.66	364.71
L1CAM	480	443.096	94.23	415.186	54.46	210.8	1016.9	806.1	1.71	5.52	4.30
AnxA2	480	4872.02	1420.40	4445.05	670.08	2711.4	17269.9	14558.5	3.63	22.145	64.83
GzmA	480	3227.45	919.83	3003.39	434.31	1650.4	12303.7	10653.3	4.03	30.65	41.98
TNFSF4	480	624.67	1287.56	520.13	55.34	348.2	27772	27423.8	19.70	409.29	58.76
GFRA1	480	2184.81	501.55	2137.60	201.63	1294.1	9214	7919.9	6.26	79.84	22.89
Shh	480	1935.51	667.40	1714.73	381.49	869.4	5264.3	4394.9	1.76	4.11	30.46
NKG2D ligand 2	480	817.430	2625.18	587.809	69.37	442.4	55713.7	55271.3	19.40	397.16	119.82
NOV	480	1653.96	441.95	1620.36	138.44	1061.2	8419.6	7358.4	9.73	128.24	20.172
FUT3	480	767.29	576.83	680.58	133.73	411.2	11626.1	11214.9	14.78	264.16	26.32
CD5L	480	2099.58	1578.77	1578.7	513.79	568.4	12931.2	12362.8	3.43	14.42	72.06
Granulins	480	29752.56	5108.33	28000.05	2417.00	18166.3	65988.6	47822.3	2.09	8.14	233.16
RAC3	480	4415.65	792.99	4544.67	431.73	2973.4	14009.5	11036.1	4.79	50.00	36.19
LY9	480	6116.85	1508.32	5589.55	934.03	2614.6	13525	10910.4	1.28	2.17	68.84
PTP1B	480	9148.64	3592.51	8977.20	1832.34	1171.5	33384.1	32212.6	0.28	4.68	163.97
RET	480	2415.88	1168.61	2212.22	477.36	829.5	21078.2	20248.7	9.09	135.04	53.33
BOC	480	1839.67	444.29	1745.91	279.04	942	6959.5	6017.5	3.66	35.76	20.27
GDF11/8	480	1042.54	230.90	1009.42	158.93	518.7	2213.3	1694.6	1.38	3.56	10.53

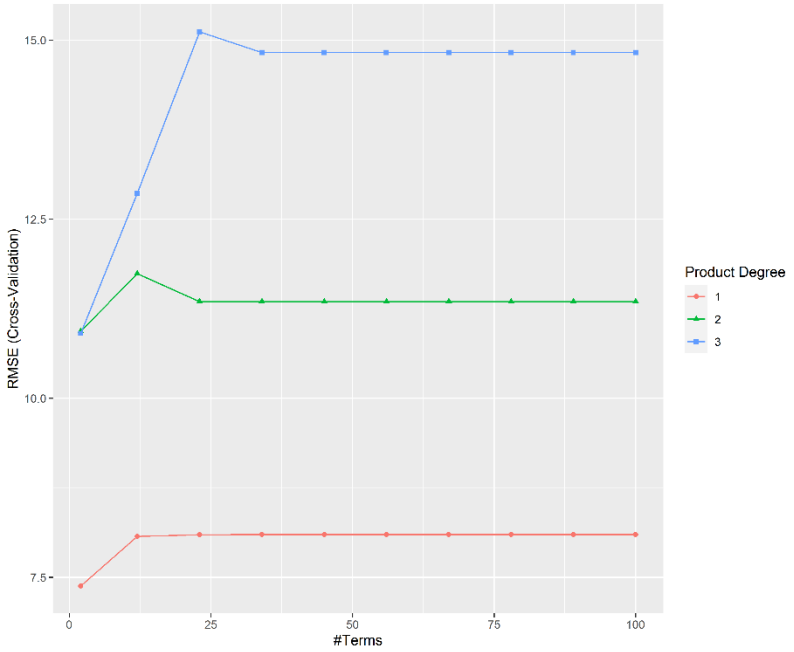
Correlation heat map of biomarkers and clinical markers

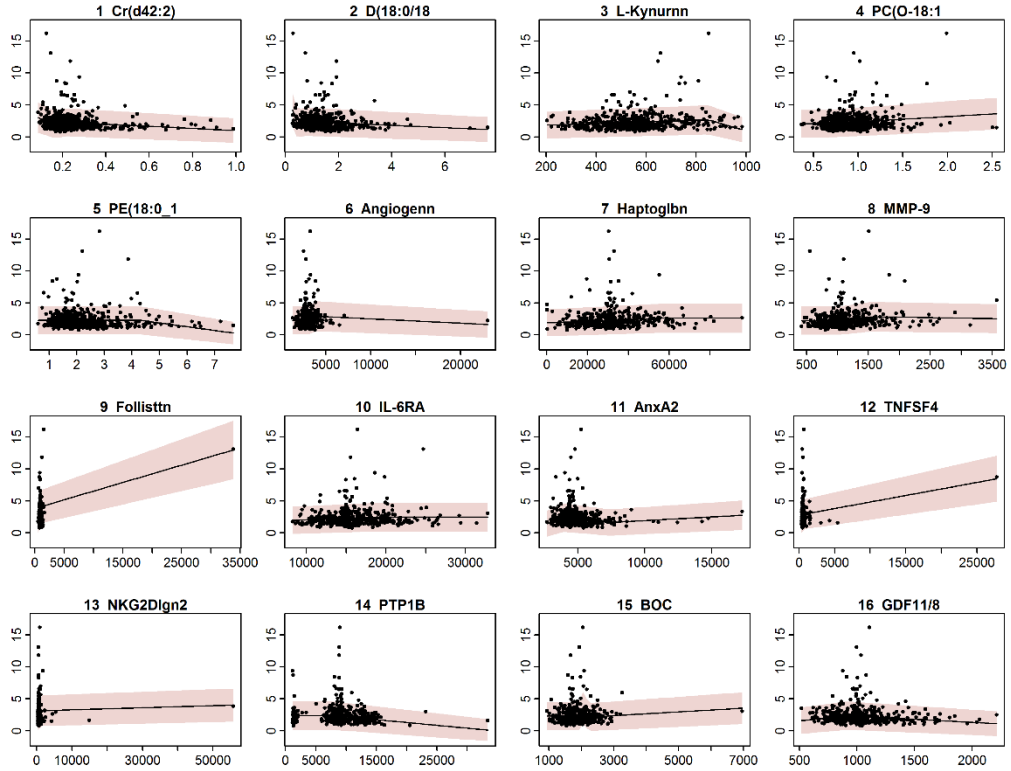


Visualizing complex relationships of biomarkers with NLR using Partial dependence plots

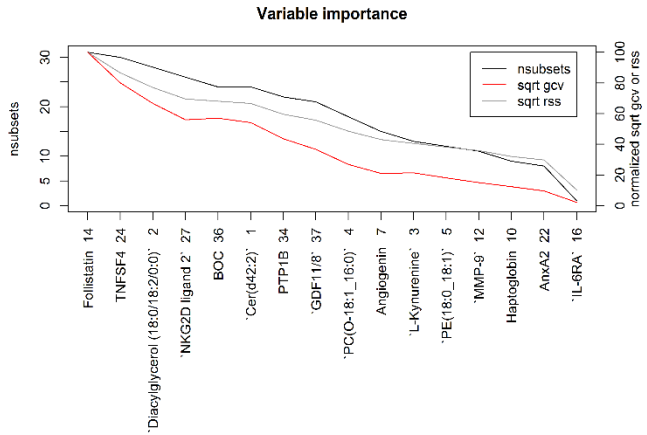


Best Interaction term for the model (5 fold CV)

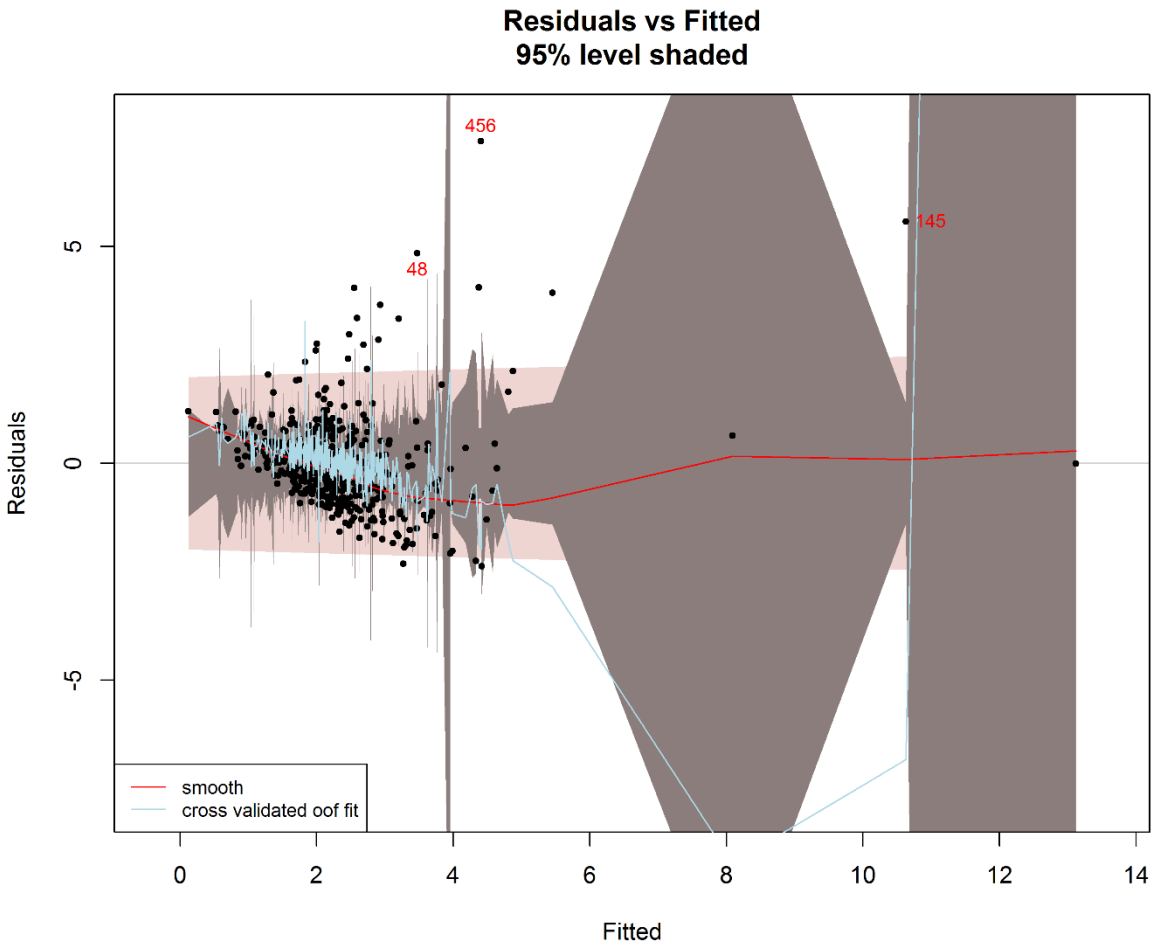


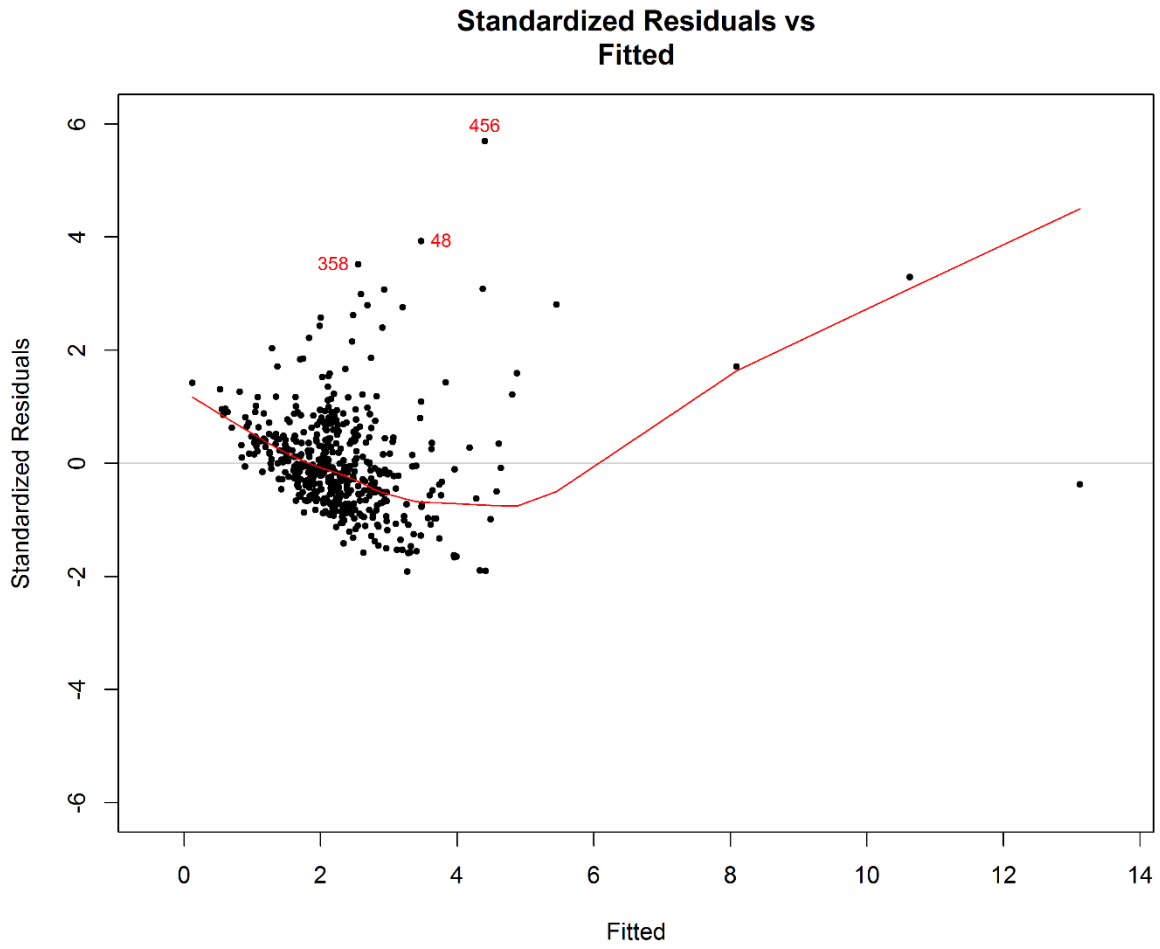


Estimate variable importance from 3-fold cross validation model using Generalized cross validation (GCV)

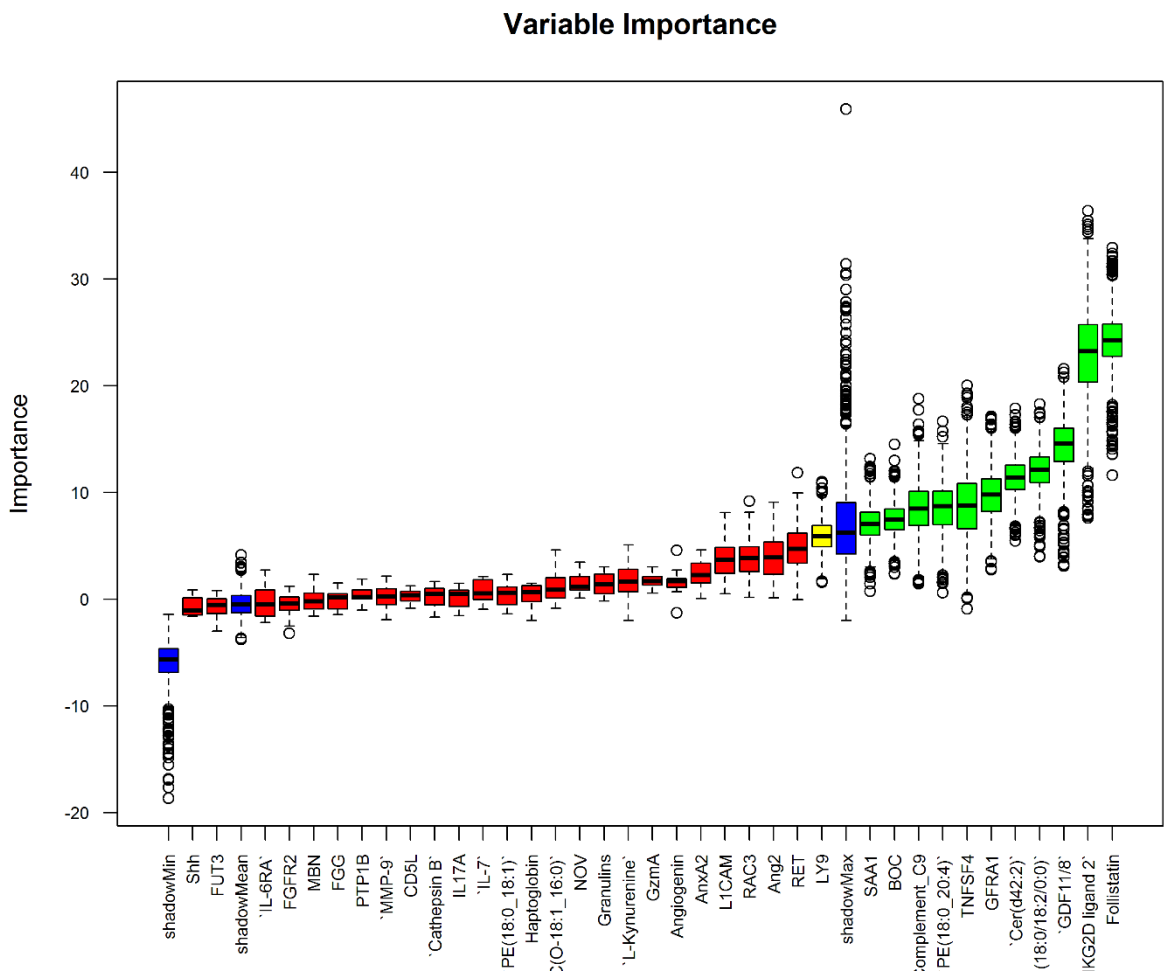


Plot showing model diagnostics

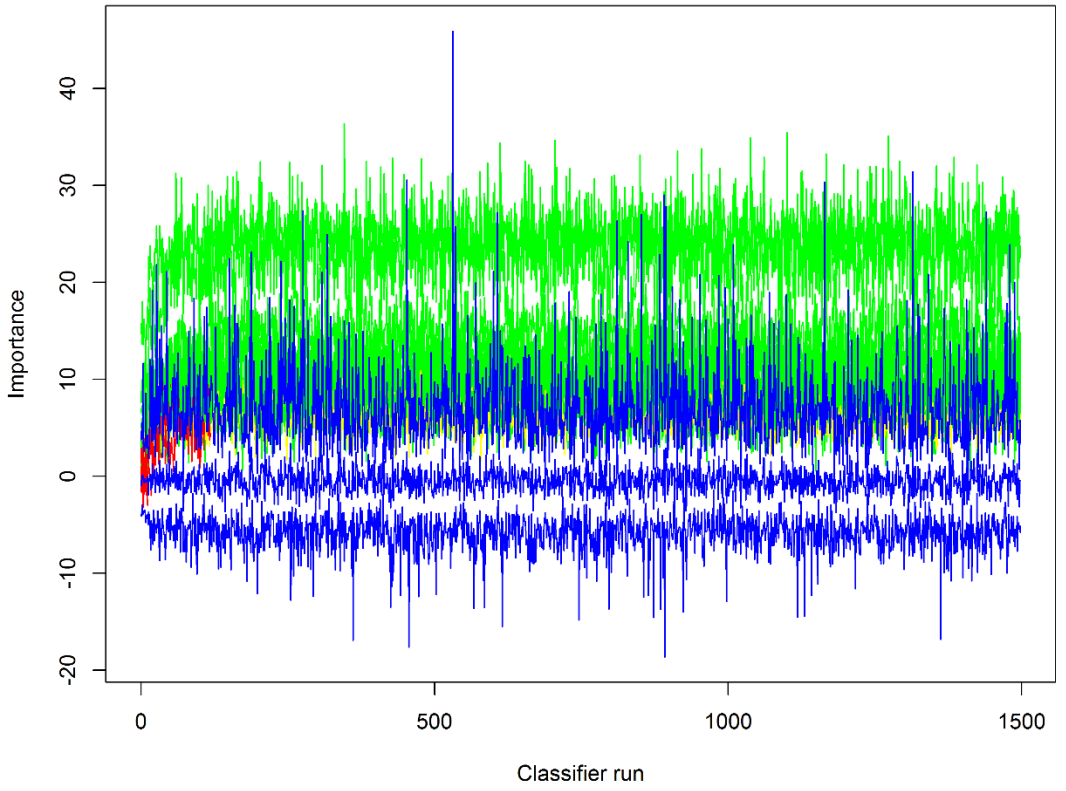




Random permutation procedure using random forest algorithm



Matplot of attribute importance over run

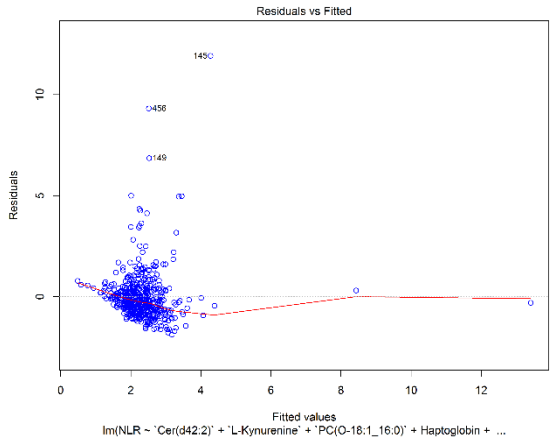


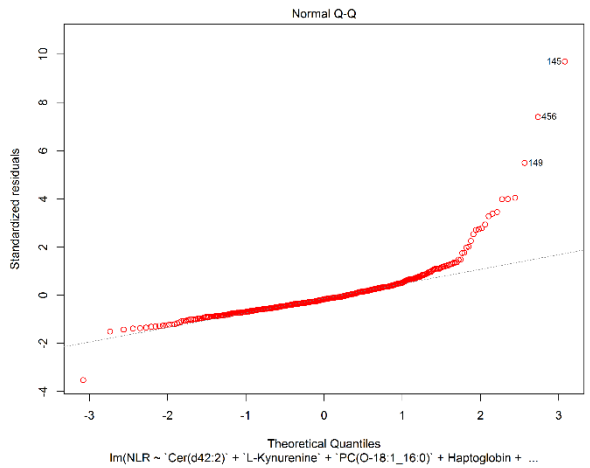
Important markers selected using Stepwise linear regression (AIC based)¥

Biomarker	Df	Sum of Sq	RSS	AIC
<none>			731.27	232.08
DAG(18:0/18:2/0:0)	1	3.497	734.76	232.37
AnxA2	1	3.798	735.06	232.56
PTP1B	1	4.823	736.09	233.23
RET	1	6.537	737.80	234.35
L1CAM	1	6.995	738.26	234.65
BOC	1	7.999	739.26	235.30
Haptoglobin	1	8.383	739.65	235.55
L-Kynurenine	1	9.839	741.11	236.49
IL17A	1	10.114	741.38	236.67
MBN	1	10.253	741.52	236.76
Cer(d42:2)	1	13.340	744.61	238.75
PC(O-18:1_16:0)	1	29.659	760.93	249.16
TNFSF4	1	33.509	764.78	251.58
Follistatin	1	122.064	853.33	304.17

untransformed independent variables¥

F-statistic: 13.38, Adjusted R-squared: 0.26





Significant markers carried forward from the AIC based model selection and conditioned

	NLR		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.28	-0.69 – 1.25	0.570
Cer(d42 : 2)	-1.82	-2.74 – -0.90	**<0.001**
L-Kynurenine	0.00	0.00 – 0.00	**0.015**
PC(O-18 : 1_16 : 0)	0.99	0.55 – 1.44	**<0.001**
Haptoglobin	0.00	0.00 – 0.00	**0.021**
Follistatin	0.00	0.00 – 0.00	**<0.001**
IL17A	0.00	0.00 – 0.00	**0.012**
MBN	0.00	0.00 – 0.00	**0.016**
L1CAM	-0.00	-0.00 – -0.00	**0.016**
TNFSF4	0.00	0.00 – 0.00	**<0.001**
RET	-0.00	-0.00 – -0.00	**0.049**
BOC	0.00	0.00 – 0.00	**0.022**
Observations	480		
R^2 / R^2 adjusted	0.275 / 0.258		

GWAS Meta analysis - NLR

Aravind Lathika Rajendrakumar

Outline of data curation and analysis

- **Phenotype**

– removed values above and below 5 SD for Neutrophils and Lymphocytes
– Removed those NLR readings with ratio change more than previous value (> 2.5)
– Removed NLR readings above 5 for those individuals with single records
– Summarised NLR (all records of NLR summarised to median value after quality checks)

- **Quality checks (No data available for other diseases in MDRF)**

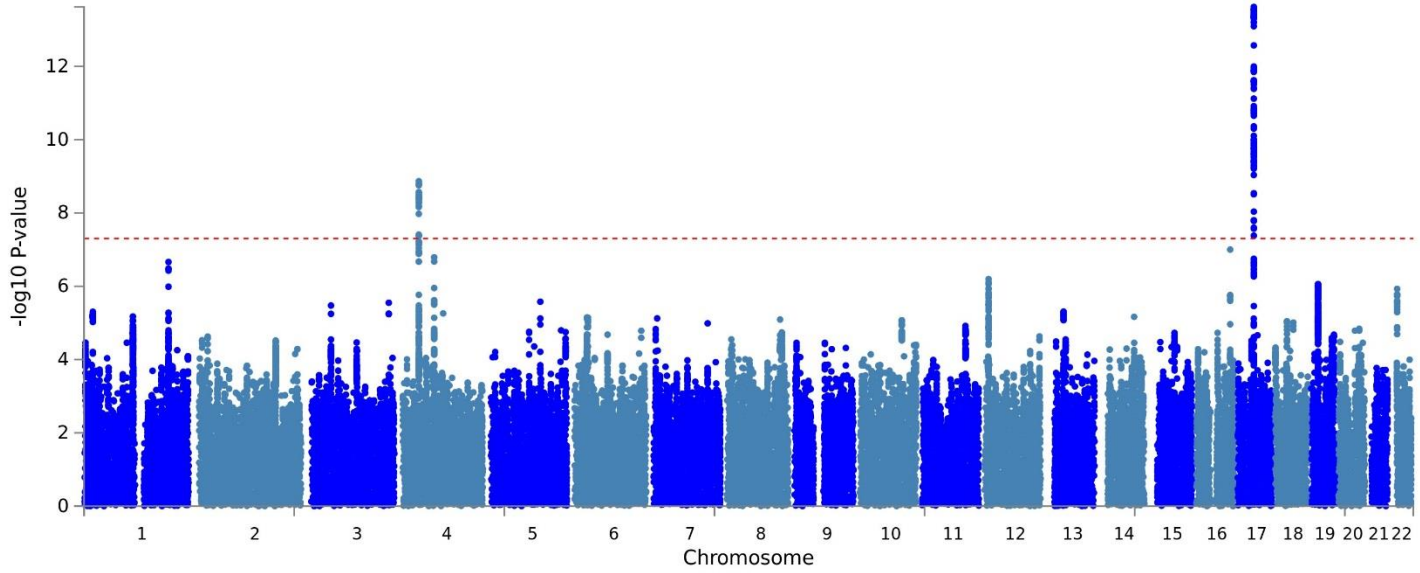
  -Excluded NLR readings after diagnosis of cancer
  -Excluded NLR readings after diagnosis of infections > 31 days
  - Retained the first NLR reading if multiple reading reported in 28 days

- All the Cohorts were adjusted for PCs, Age, Sex and DM status
- SNP test was used for GWAS analysis
- GWAMA was used for fixed effects meta analysis – GoDARTS + GoSHARE

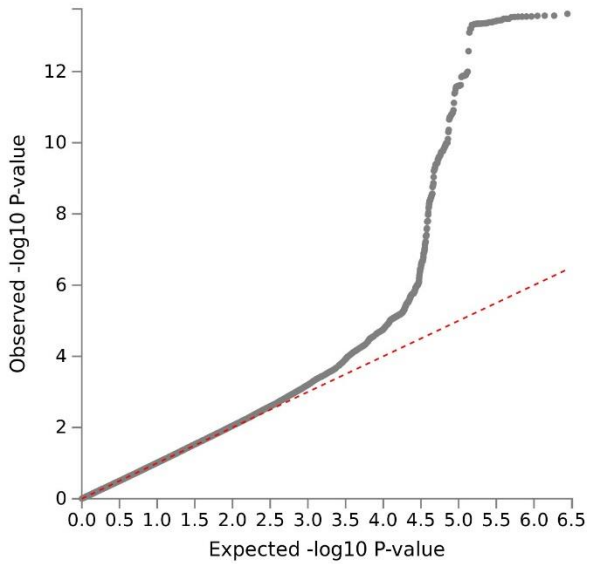
Overall Cohort Description (n= 21,153)

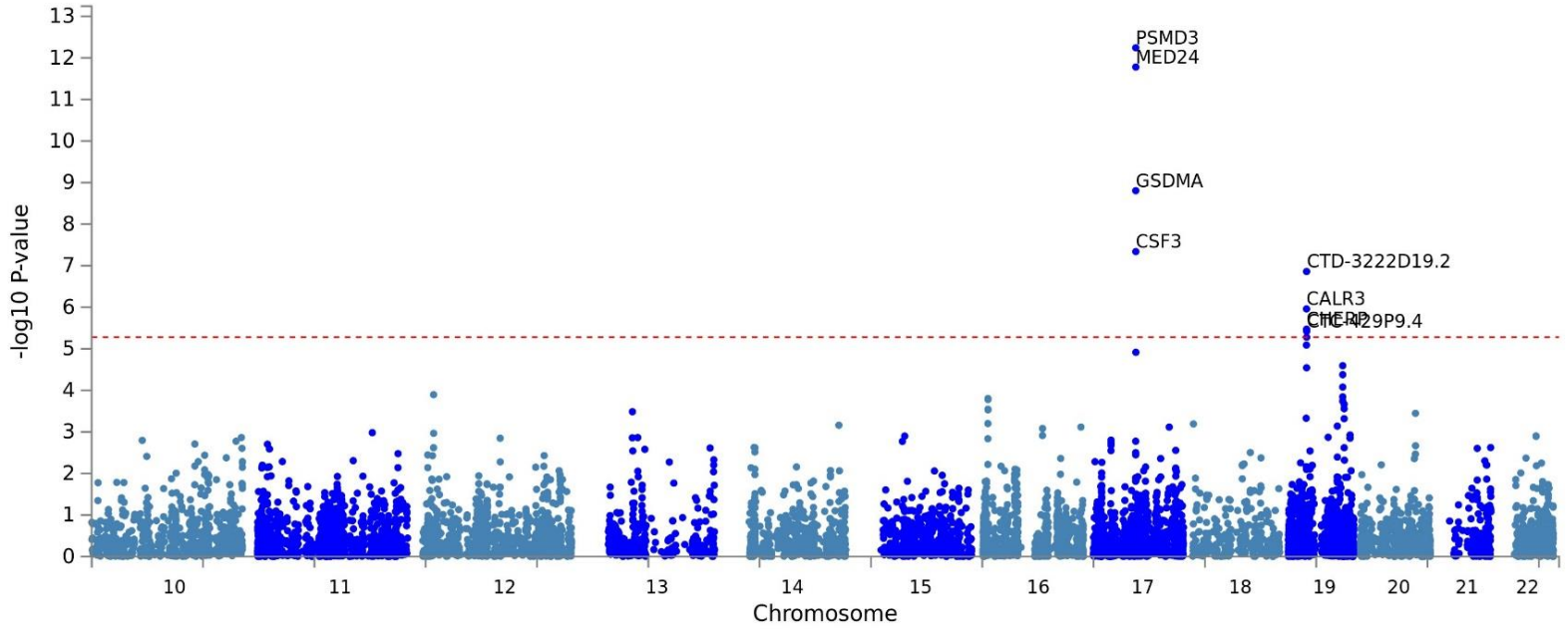
Cohort	Sample Size	Median NLR (IQR)	Median Age (IQR)	Sex (M%)	DM (%)	GIF (λ)
GoDARTS						
Affymetrix	3,754	2.30 (1.83-2.93)	74.75 (67.01-81.43)	2001(53.3%)	3651(97.2%)	1.02
Illumina	3,549	2.29 (1.75-2.94)	74.19 (65.43- 81.42)	2035(57.3%)	3417(96.3%)	1.04
Broad	1,834	2.15 (1.68-2.70)	70.12 (59.03-79.42)	924(50.4%)	951(51.8%)	0.99
GoSHARE	6,446	2.14 (1.70-2.73)	66.49 (56.28-75.30)	3581 (52.8%)	4891(72.1%)	1.01
GDF2	5,570	2.04 (1.62-2.71)	69.39 (58.72 -79.75)	2727(48.9%)	433 (7.7%)	1.02
GoSHARE +GoSHARE + GDF2	21,153	-	-	--	--	1.09

GWAS of NLR in the Scottish Cohort

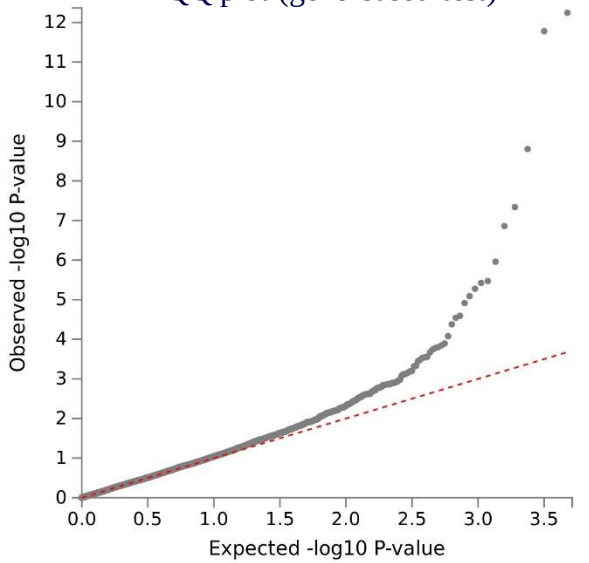


QQ plot (GWAS summary statistics)



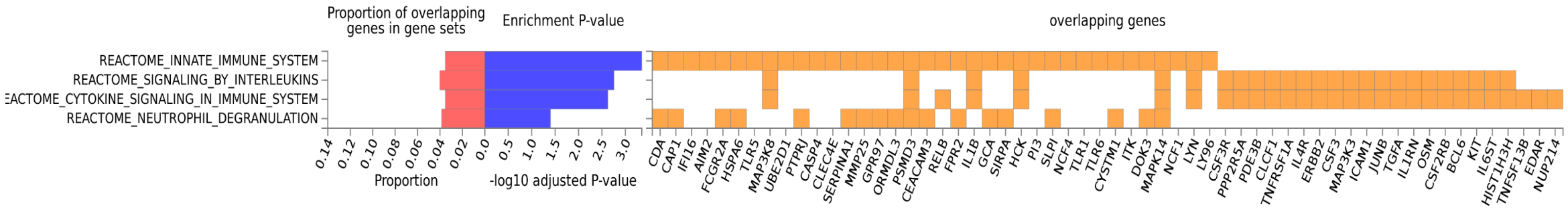


QQ plot (gene-based test)



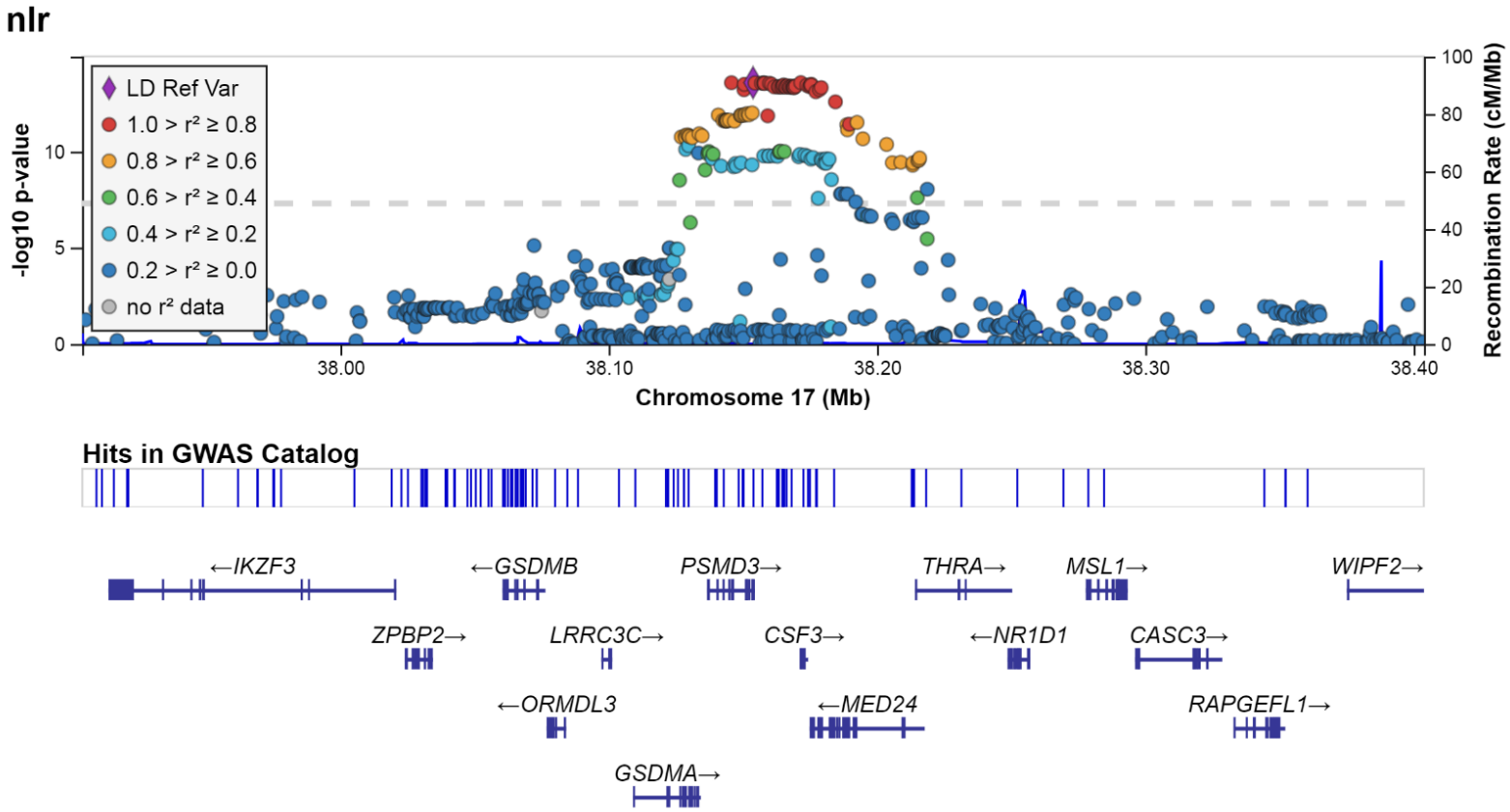
GWAS Catalogue reported genes

Reactome Pathways

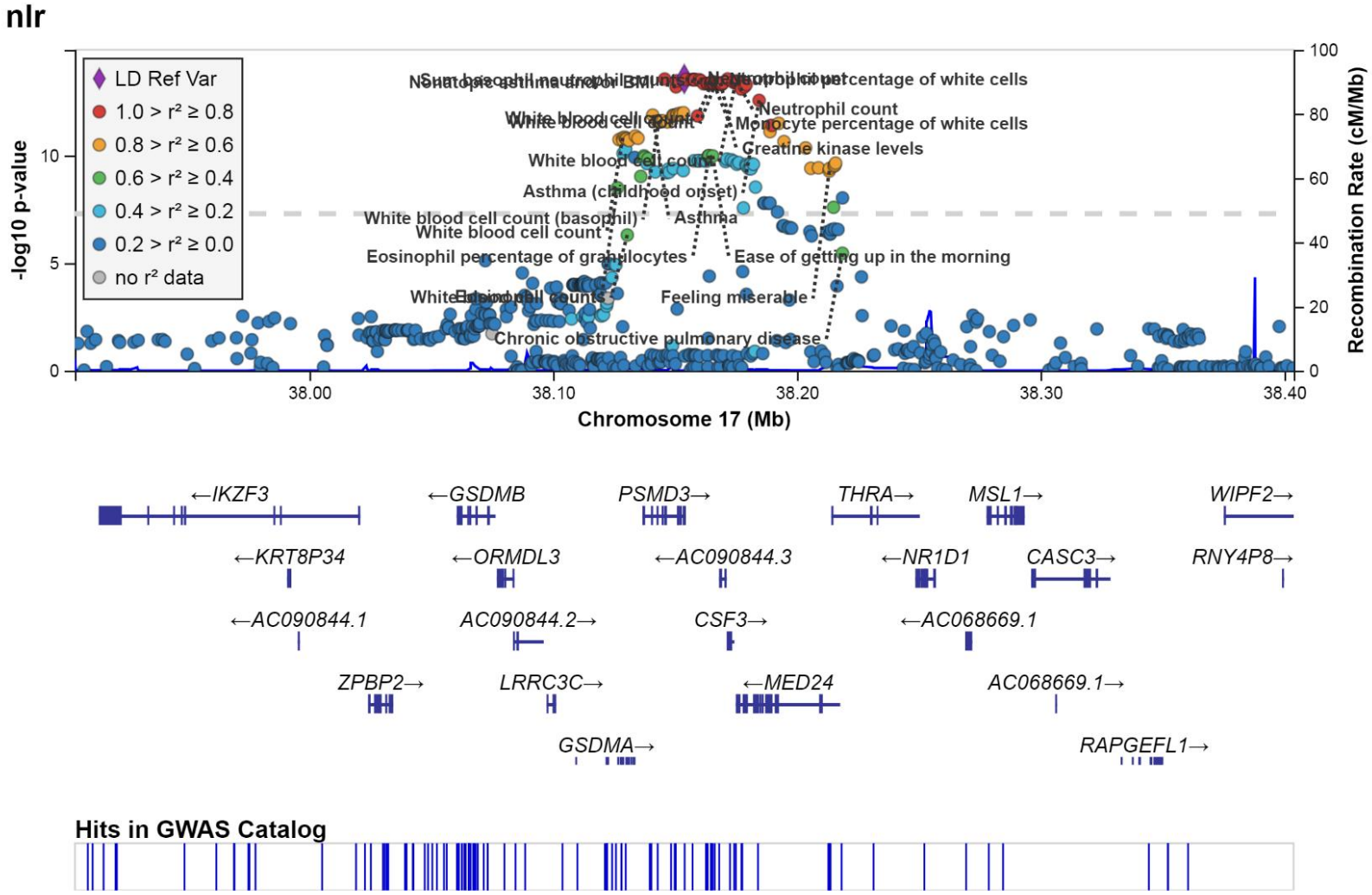


Source: https://fuma.ctglab.nl/

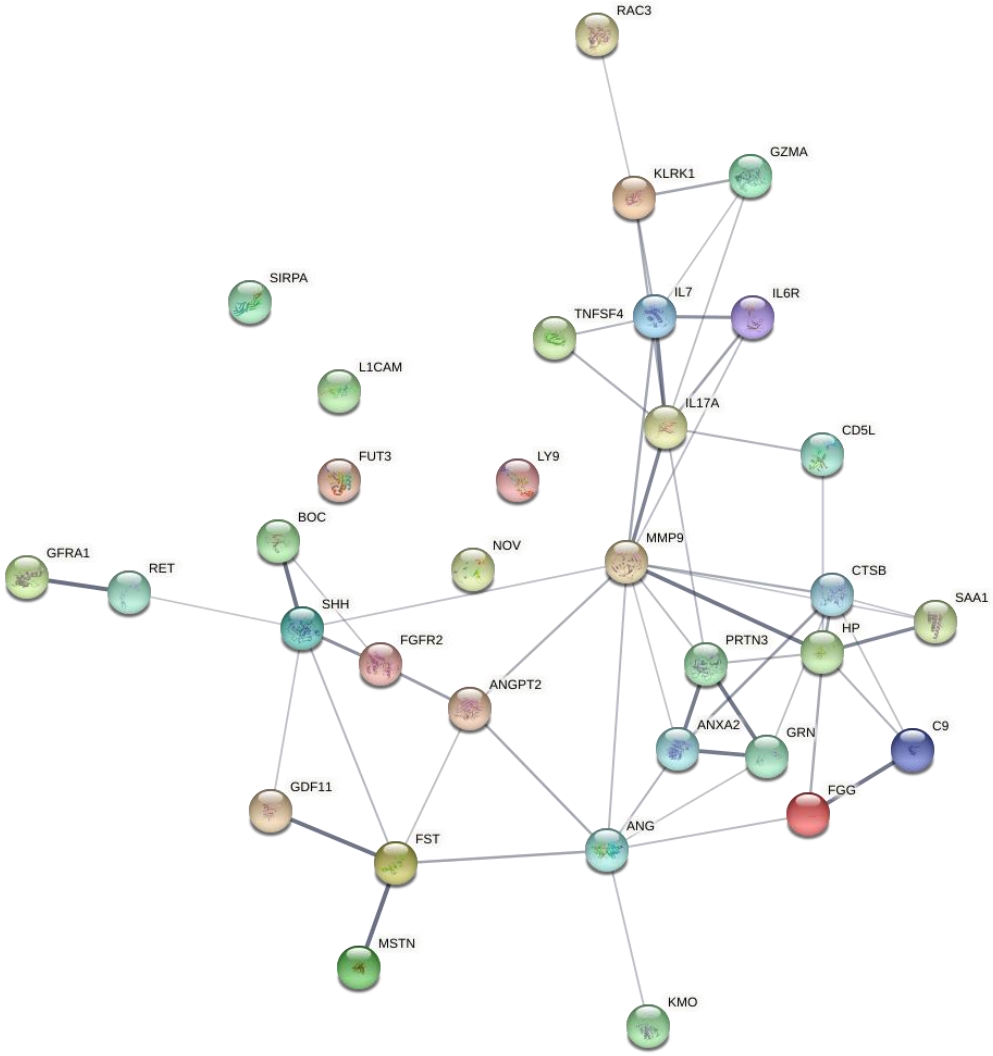
Locus Zoom Plot of Chromosome 17



Locus Zoom Plot of Chromosome 17

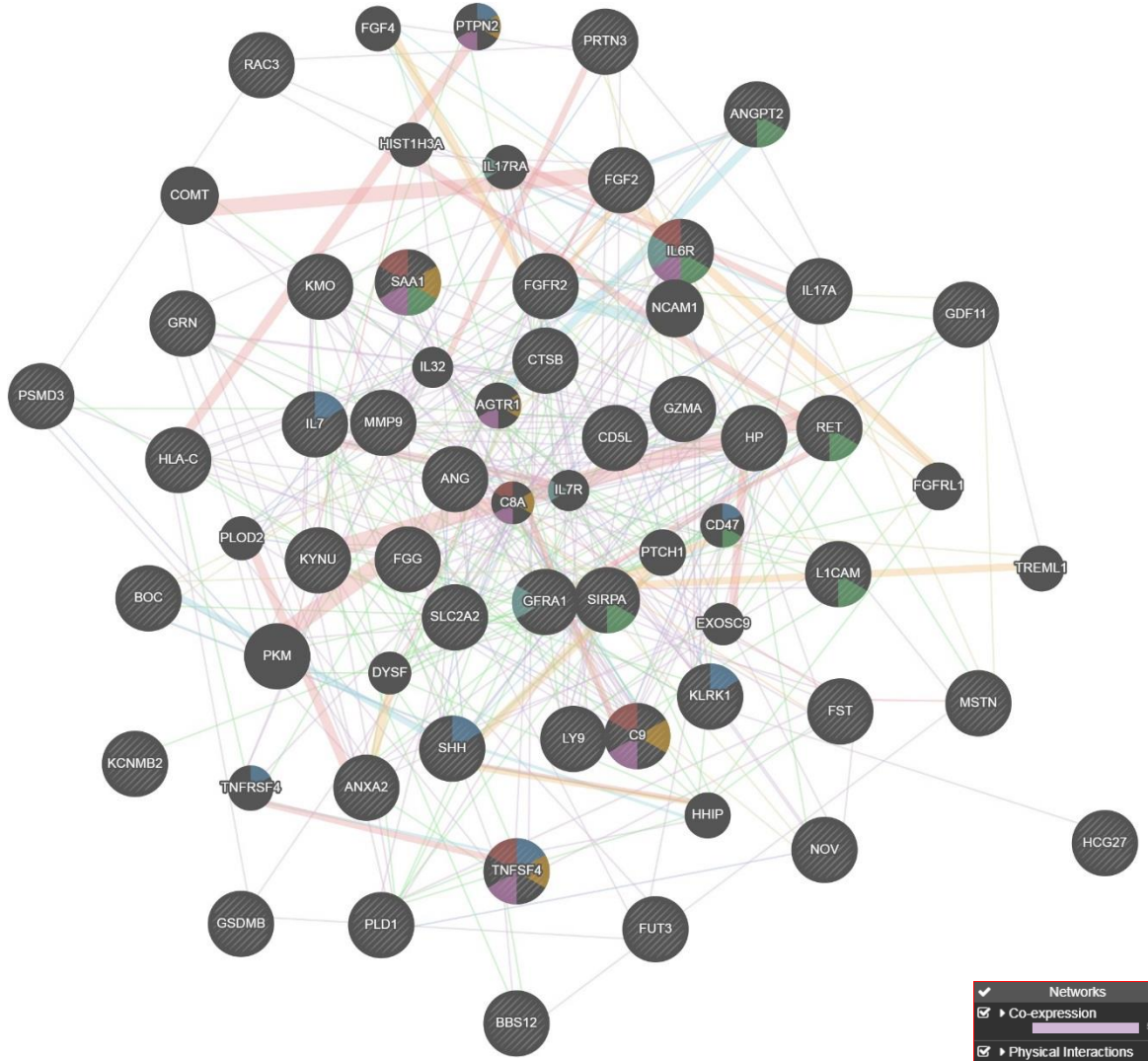


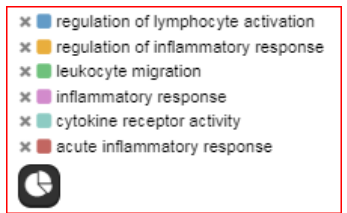
Protein-Protein Interaction Network of biomarkers selected from LASSO Regression

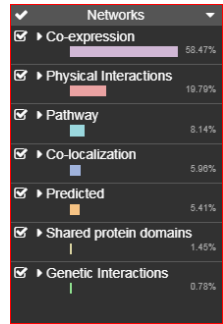


number of nodes: 33
number of edges: 55
average node degree: 3.33
avg. local clustering coefficient: 0.48
expected number of edges: 11
PPI enrichment p-value:< 1.0e-16

Gene-Gene Interaction Network for biomarker variables with NLR and its sub components

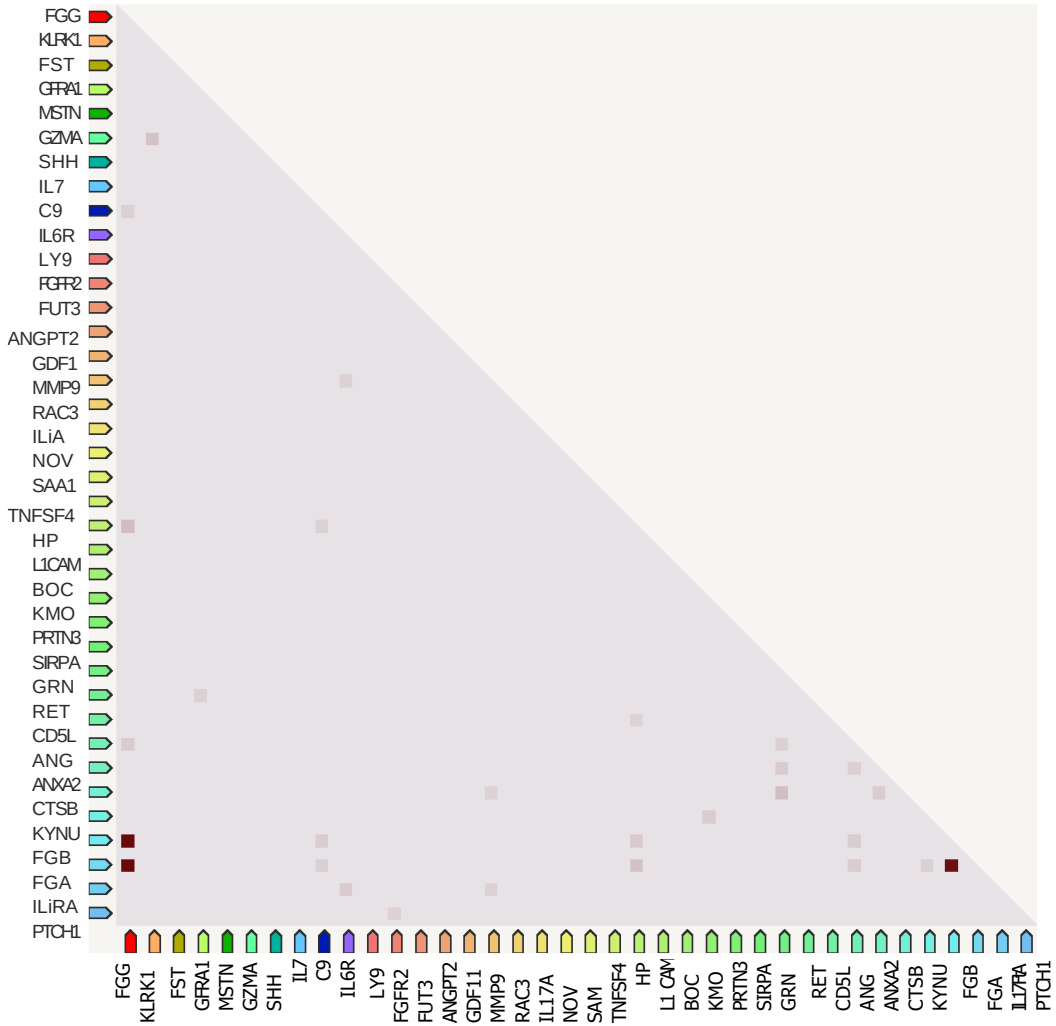




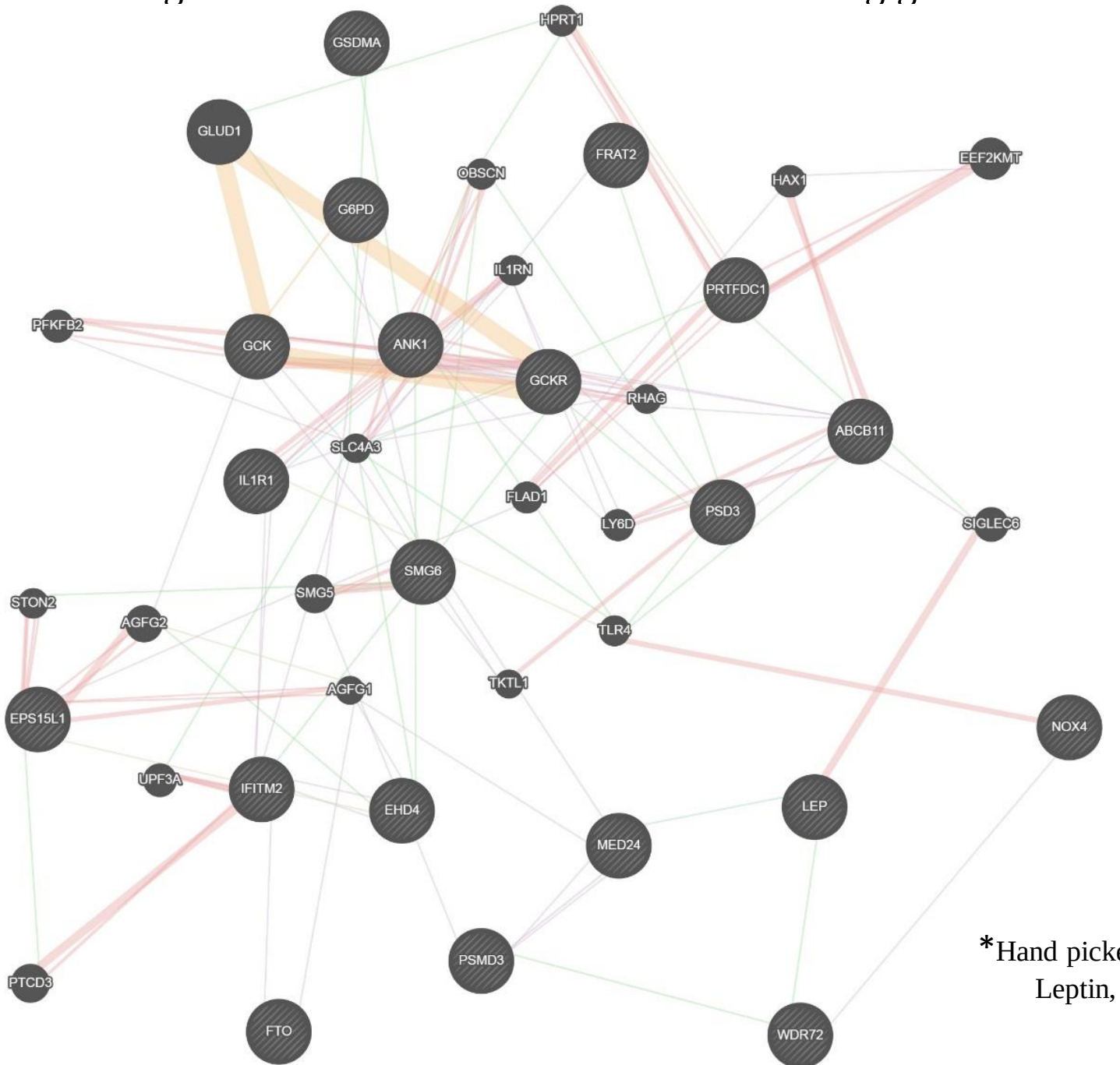


Gene Expression

observed Co-expression in Homo sapiens:

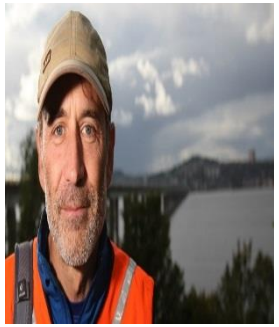


Understanding the inflammation at molecular level using gene network*



*Hand picked established genes for HbA1c, Leptin, DR, BMI, Neutrophils and Lymphocytes

Acknowledgement



- All Members of INSPIRED Research group.
- MDRF team
- Health Informatics Centre (HIC)- University of Dundee
- National Institute for Health Research (NIHR)

This research was commissioned by the National Institute for Health Research using Official Development Assistance (ODA) funding [INSPIRED 16/136/102].

"*The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.*"

FUNDED BY
NIHR | National Institute for Health Research

Thank You

Questions?